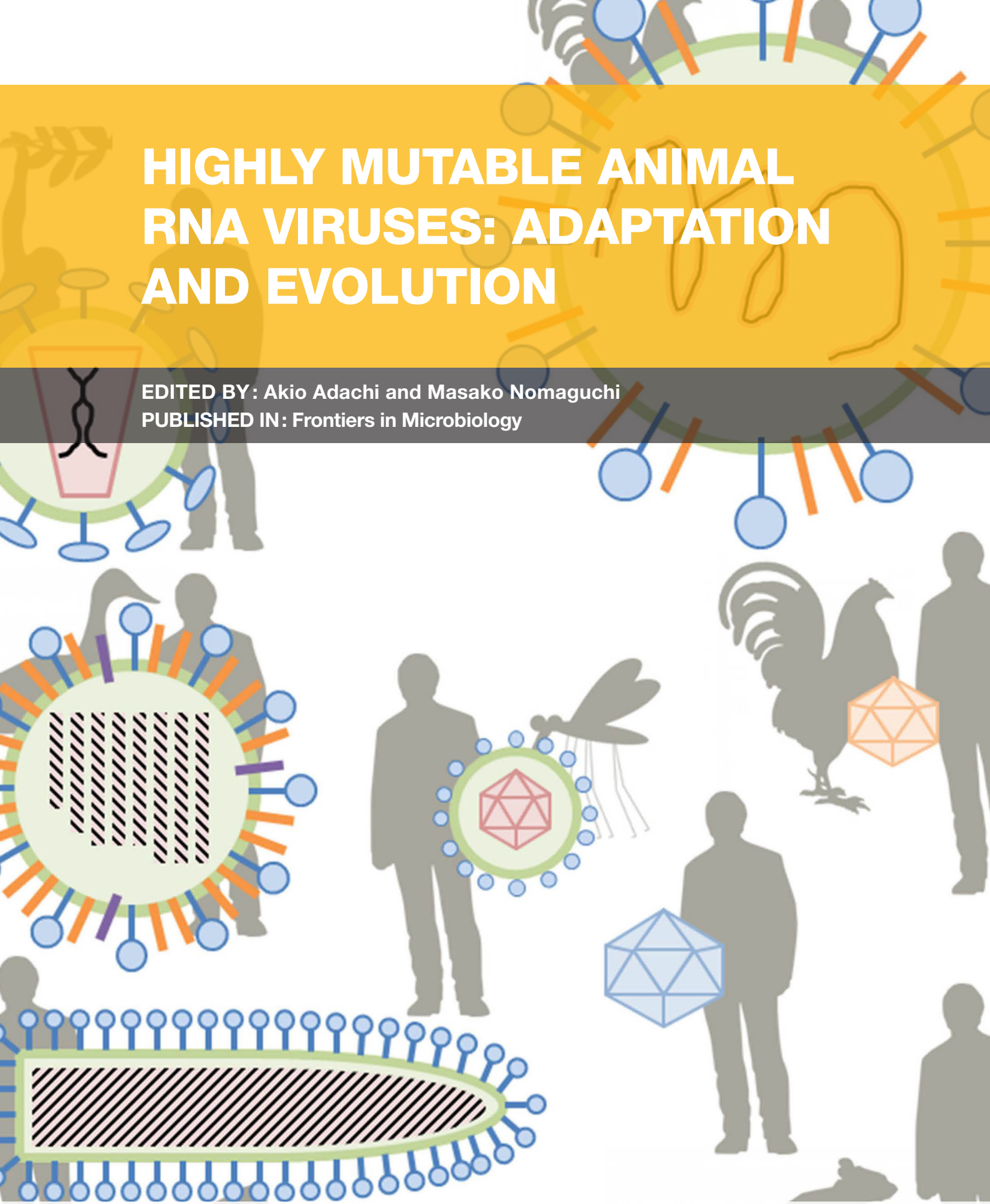
HIGHLY MUTABLE ANIMAL RNA VIRUSES: ADAPTATION AND EVOLUTION

EDITED BY: Akio Adachi and Masako Nomaguchi
PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics



frontiers

Frontiers Copyright Statement

© Copyright 2007-2017 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88945-348-1
DOI 10.3389/ 978-2-88945-348-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

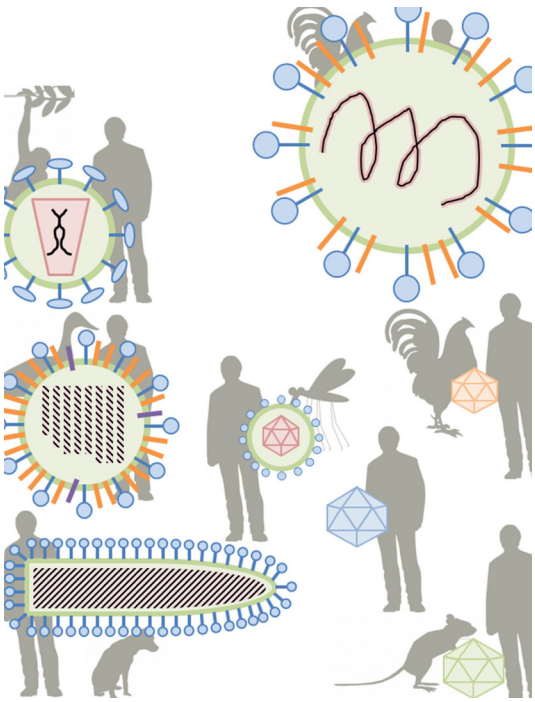
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

HIGHLY MUTABLE ANIMAL RNA VIRUSES: ADAPTATION AND EVOLUTION

Topic Editors:
Akio Adachi, Tokushima University and Kansai Medical University, Japan
Masako Nomaguchi, Tokushima University, Japan



Cartoonish virion representation of various animal RNA viruses in this Research Topic. Virion-morphology of lentivirus, orthomyxovirus, paramyxovirus, flavivirus, rhabdovirus, calicivirus, picornavirus, and picobirnavirus is illustrated against its natural hosts (silhouettes from Clipart free-images). Virions are enlarged proportionally to their real sizes.

Image drawn by Masako Nomaguchi (Tokushima University, Tokushima, Japan), a host-editor of this Research Topic.

Viruses are widely present in nature, and numerous viral species with a variety of unique characteristics have been identified so far. Even now, new emerging or re-emerging viruses are being found or re-found as novel viral classes or as quasi-species. Indeed, viruses are everywhere. Of note, viruses are pivotal as targets and tools of basic and applied sciences. On one hand, portions of the viruses are infectious for animals including humans, and cause various diseases in infected hosts by distinct mechanisms and at a different level of severity. While many of viruses

are known to co-exist quietly with their hosts, pathogenic viruses certainly affect and threaten our society as well as individuals to provoke serious medical or economic attention. We should act against certain dreadful and highly infectious viruses as a global problem.

Animal RNA viruses can readily mutate to adapt themselves in their hostile environments for their survival. Resultant viruses may sometimes show essentially altered phenotypes from the original parental strains. This fundamental and general property of animal RNA viruses represents major extensive issues of scientific, medical, and/or economic importance. In this Research Topic, we have focused on the high mutability of animal RNA viruses, and selected relevant articles on animal viruses of broad-ranges such as primate lentiviruses, influenza viruses, paramyxoviruses, flaviviruses, rabies virus, norovirus, picornaviruses, and picobirnavirus. Each article has taken up intriguing aspects of the subject viruses. We are sure that readers acquire important information on virus mutation, adaptation, diversification, and evolution, and hope that researchers in the field related to virology gain some solid hints from the reported articles for further virological and /or medical studies. Finally, we thank all the contributing researchers in this Research Topic, entitled "Highly Mutable Animal RNA Viruses: Adaptation and Evolution", for their elegant and interesting works.

Citation: Adachi, A., Nomaguchi, M., eds. (2017). Highly Mutable Animal RNA Viruses: Adaptation and Evolution. Lausanne: Frontiers Media. doi: 10.3389/ 978-2-88945-348-1

Table of Contents

08 ***Editorial: Highly Mutable Animal RNA Viruses: Adaptation and Evolution***
Masako Nomaguchi and Akio Adachi

Chapter 1: Primate lentiviruses

10 ***Primate Lentiviruses Modulate NF-κB Activity by Multiple Mechanisms to Fine-Tune Viral and Cellular Gene Expression***
Elena Heusinger and Frank Kirchhoff

22 ***HIV-1 Tat and Viral Latency: What We Can Learn from Naturally Occurring Sequence Variations***
Doreen Kamori and Takamasa Ueno

29 ***Impact of Genetic Variations in HIV-1 Tat on LTR-Mediated Transcription via TAR RNA Interaction***
Larance Ronsard, Nilanjana Ganguli, Vivek K. Singh, Kumaravel Mohankumar, Tripti Rai, Subhashree Sridharan, Sankar Pajaniradje, Binod Kumar, Devesh Rai, Suhnrita Chaudhuri, Mohane S. Coumar, Vishnampettai G. Ramachandran and Akhil C. Banerjea

42 ***Unusual Fusion Proteins of HIV-1***
Simon Langer and Daniel Sauter

53 ***Driving HIV-1 into a Vulnerable Corner by Taking Advantage of Viral Adaptation and Evolution***
Shigeyoshi Harada and Kazuhisa Yoshimura

61 ***Commentary: MARCH8 Inhibits HIV-1 Infection by Reducing Virion Incorporation of Envelope Glycoproteins***
Mikako Fujita

64 ***APOBEC3G-Mediated G-to-A Hypermutation of the HIV-1 Genome: The Missing Link in Antiviral Molecular Mechanisms***
Ayaka Okada and Yasumasa Iwatani

72 ***Expression Profiles of Vpx/Vpr Proteins Are Co-related with the Primate Lentiviral Lineage***
Yosuke Sakai, Ariko Miyake, Naoya Doi, Hikari Sasada, Yasuyuki Miyazaki, Akio Adachi and Masako Nomaguchi

83 ***Phylogenetic Insights into the Functional Relationship between Primate Lentiviral Reverse Transcriptase and Accessory Proteins Vpx/Vpr***
Yosuke Sakai, Naoya Doi, Yasuyuki Miyazaki, Akio Adachi and Masako Nomaguchi

88 ***Comparison of Biochemical Properties of HIV-1 and HIV-2 Capsid Proteins***
Yasuyuki Miyazaki, Ariko Miyake, Noya Doi, Takaaki Koma, Tsuneo Uchiyama, Akio Adachi and Masako Nomaguchi

Chapter 2: Influenza viruses

99 Pathogenicity, Transmission and Antigenic Variation of H5N1 Highly Pathogenic Avian Influenza Viruses

Peirong Jiao, Hui Song, Xiaoke Liu, Yafen Song, Jin Cui, Siyu Wu, Jiaqi Ye, Nanan Qu, Tiemin Zhang and Ming Liao

107 Immune Responses of Chickens Infected with Wild Bird-Origin H5N6 Avian Influenza Virus

Shimin Gao, Yinfeng Kang, Runyu Yuan, Haili Ma, Bin Xiang, Zhaoxiong Wang, Xu Dai, Fumin Wang, Jiajie Xiao, Ming Liao and Tao Ren

116 Pathogenesis and Phylogenetic Analyses of Two Avian Influenza H7N1 Viruses Isolated from Wild Birds

Hongmei Jin, Deli Wang, Jing Sun, Yanfang Cui, Guang Chen, Xiaolin Zhang, Jiajie Zhang, Xiang Li, Hongliang Chai, Yuwei Gao, Yanbing Li and Yuping Hua

133 Two Genetically Similar H9N2 Influenza A Viruses Show Different Pathogenicity in Mice

Qingtao Liu, Yuzhuo Liu, Jing Yang, Xinmei Huang, Kaikai Han, Dongmin Zhao, Keran Bi and Yin Li

146 Influenza A Viruses Replicate Productively in Mouse Mastocytoma Cells (P815) and Trigger Pro-inflammatory Cytokine and Chemokine Production through TLR3 Signaling Pathway

Di Meng, Caiyun Huo, Ming Wang, Jin Xiao, Bo Liu, Tangting Wei, Hong Dong, Guozhong Zhang, Yanxin Hu and Lunquan Sun

159 Tracking the Evolution of Polymerase Genes of Influenza A Viruses during Interspecies Transmission between Avian and Swine Hosts

Nipawit Karnbunchob, Ryosuke Omori, Heidi L. Tessmer and Kimihito Ito

171 A Novel H1N2 Influenza Virus Related to the Classical and Human Influenza Viruses from Pigs in Southern China

Yafen Song, Xiaowei Wu, Nianchen Wang, Guowen Ouyang, Nannan Qu, Jin Cui, Yan Qi, Ming Liao and Peirong Jiao

184 Continuing Reassortant of H5N6 Subtype Highly Pathogenic Avian Influenza Virus in Guangdong

Runyu Yuan, Zheng Wang, Yinfeng Kang, Jie Wu, Lirong Zou, Lijun Liang, Yingchao Song, Xin Zhang, Hanzhong Ni, Jinyan Lin and Changwen Ke

199 Novel H7N2 and H5N6 Avian Influenza A Viruses in Sentinel Chickens: A Sentinel Chicken Surveillance Study

Teng Zhao, Yan-Hua Qian, Shan-Hui Chen, Guo-Lin Wang, Meng-Na Wu, Yong Huang, Guang-Yuan Ma, Li-Qun Fang, Gregory C. Gray, Bing Lu, Yi-Gang Tong, Mai-Juan Ma and Wu-Chun Cao

208 Molecular Dynamics Simulation of the Influenza A(H3N2) Hemagglutinin Trimer Reveals the Structural Basis for Adaptive Evolution of the Recent Epidemic Clade 3C.2a

Masaru Yokoyama, Seiichiro Fujisaki, Masayuki Shirakura, Shinji Watanabe, Takato Odagiri, Kimito Ito and Hironori Sato

Chapter 3: Paramyxoviruses

218 Detection of Inter-Lineage Natural Recombination in Avian Paramyxovirus Serotype 1 Using Simplified Deep Sequencing Platform

Dilan A. Satharasinghe, Kavitha Murulitharan, Sheau W. Tan, Swee K. Yeap, Muhammad Munir, Aini Ideris and Abdul R. Omar

232 ***Phylogenetic and Pathotypic Characterization of Newcastle Disease Viruses Circulating in South China and Transmission in Different Birds***
Yinfeng Kang, Bin Xiang, Runyu Yuan, Xiaqiong Zhao, Minsha Feng, Pei Gao, Yanling Li, Yulian Li, Zhangyong Ning and Tao Ren

245 ***Infection of Goose with Genotype VIId Newcastle Disease Virus of Goose Origin Elicits Strong Immune Responses at Early Stage***
Qianqian Xu, Yuqiu Chen, Wenjun Zhao, Tingting Zhang, Chenggang Liu, Tianming Qi, Zongxi Han, Yuhao Shao, Deying Ma and Shengwang Liu

258 ***180-Nucleotide Duplication in the G Gene of* Human metapneumovirus *A2b Subgroup Strains Circulating in Yokohama City, Japan, since 2014***
Miwako Saikusa, Chiharu Kawakami, Naganori Nao, Makoto Takeda, Shuzo Usuku, Tadayoshi Sasao, Kimiko Nishimoto and Takahiro Toyozawa

269 ***Evolution and Transmission of Respiratory Syncytial Group A (RSV-A) Viruses in Guangdong, China 2008–2015***
Lirong Zou, Lina Yi, Jie Wu, Yingchao Song, Guofeng Huang, Xin Zhang, Lijun Liang, Hanzhong Ni, Oliver G. Pybus, Changwen Ke and Jing Lu

Chapter 4: Flaviviruses

280 ***Genetic Diversity and Positive Selection Analysis of Classical Swine Fever Virus Envelope Protein Gene E2 in East China under C-Strain Vaccination***
Dongfang Hu, Lin Lv, Jinyuan Gu, Tongyu Chen, Yihong Xiao and Sidang Liu

290 ***GRIM-19 Restricts HCV Replication by Attenuating Intracellular Lipid Accumulation***
Jung-Hee Kim, Pil S. Sung, Eun B. Lee, Wonhee Hur, Dong J. Park, Eui-Cheol Shin, Marc P. Windisch and Seung K. Yoon

304 ***Zika Virus: the Latest Newcomer***
Juan-Carlos Saiz, Ángela Vázquez-Calvo, Ana B. Blázquez, Teresa Merino-Ramos, Estela Escribano-Romero and Miguel A. Martín-Acebes

323 ***Development of a Zika Virus Infection Model in Cynomolgus Macaques***
Fusataka Koide, Scott Goebel, Beth Snyder, Kevin B. Walters, Alison Gast, Kimberly Hagelin, Raj Kalkeri and Jonathan Rayner

Chapter 5: Other viruses

5-1: Rabies virus

331 ***Systems Biomedicine of Rabies Delineates the Affected Signaling Pathways***
Sadegh Azimzadeh Jamalkandi, Sayed-Hamidreza Mozhgani, Hamid Gholami Pourbadie, Mehdi Mirzaie, Farshid Noorbakhsh, Behrouz Vaziri, Alireza Gholami, Naser Ansari-Pour and Mohieddin Jafari

353 ***Phenotypic Consequences* In vivo *and* In vitro *of Rearranging the P Gene of RABV HEP-Flury***
Mingzhu Mei, Teng Long, Qiong Zhang, Jing Zhao, Qin Tian, Jiaojiao Peng, Jun Luo, Yifei Wang, Yingyi Lin and Xiaofeng Guo

5-2: Norovirus

365 ***Viral Population Changes during Murine Norovirus Propagation in RAW 264.7 Cells***
Takuya Kitamoto, Reiko Takai-Todaka, Akiko Kato, Kumiko Kanamori, Hirotaka Takagi, Kazuhiro Yoshida, Kazuhiko Katayama and Akira Nakanishi

376 Evolutionary Constraints on the Norovirus Pandemic Variant GII.4_2006b over the Five-Year Persistence in Japan
Hironori Sato, Masaru Yokoyama, Hiromi Nakamura, Tomoichiro Oka, Kazuhiko Katayama, Naokazu Takeda, Mamoru Noda, Tomoyuki Tanaka and Kazushi Motomura

5-3: Picornaviruses

385 Whole Genome Sequencing of* Enterovirus *species* C *Isolates by High-Throughput Sequencing: Development of Generic Primers
Maël Bessaud, Serge A. Sadeuh-Mba, Marie-Line Joffret, Richter Razafindratsimandresy, Patsy Polston, Romain Volle, Mala Rakoto-Andrianarivelo, Bruno Blondel, Richard Njouom and Francis Delpeyroux

395 Comparative Genetic Analyses of Human Rhinovirus C (HRV-C) Complete Genome from Malaysia
Yam Sim Khaw, Yoke Fun Chan, Faizatul Lela Jafar, Norlijah Othman and Hui Yee Chee

5-4: Picobirnavirus

408 High Diversity of Genogroup I Picobirnaviruses in Mammals
Patrick C. Y. Woo, Jade L. L. Teng, Ru Bai, Annette Y. P. Wong, Paolo Martelli, Suk-Wai Hui, Alan K. L. Tsang, Candy C. Y. Lau, Syed S. Ahmed, Cyril C. Y. Yip, Garnet K. Y. Choi, Kenneth S. M. Li, Carol S. F. Lam, Susanna K. P. Lau and Kwok-Yung Yuen

frontiers in Microbiology

EDITORIAL
published: 15 September 2017
doi: 10.3389/fmicb.2017.01785

Editorial: Highly Mutable Animal RNA Viruses: Adaptation and Evolution

Masako Nomaguchi[1] *and Akio Adachi*[1, 2]*

[1] *Department of Microbiology, Tokushima University Graduate School of Medical Science, Tokushima, Tokushima, Japan,*
[2] *Department of Microbiology, Kansai Medical University, Hirakata, Osaka, Japan*

Keywords: animal viruses, RNA viruses, mutation, adaptation, evolution

Editorial on the Research Topic

Highly Mutable Animal RNA Viruses: Adaptation and Evolution

One of the most conspicuous and fundamental characteristics of animal RNA viruses is their high mutability under various circumstances. This property is a major cause for viral adaptation to changing environments, and also for medical problems/issues associated with viruses. In this Research Topic, numerous researchers have described and discussed the latest and/or momentous results on animal RNA viruses as original (27 papers), methods (1 paper), review (3 papers), mini-review (3 papers), perspective (1 paper), and general commentary (1 paper) articles. Given that each virus has its own strategy for its replication, transmission, and survival, we are justified to categorize individual articles into 8 groups by the target virus in question: (I) *Retroviridae-lentivirinae* (human and simian immunodeficiency viruses, 10 articles), (II) *Orthomyxoviridae* (influenza virus, 10 articles), (III) *Paramyxoviridae* (avian Newcastle disease virus, human metapneumovirus, and human respiratory syncytial virus, 5 articles), (IV) *Flaviviridae* (classical swine virus, hepatitis C virus, and Zika virus, 4 articles), (V) *Rhabdoviridae* (rabies virus, 2 articles), (VI) *Caliciviridae* (norovirus, 2 articles), (VII) *Picornaviridae* (enterovirus species C, and rhinovirus, 2 articles), and (VIII) *Picobirnaviridae* (picobirnavirus, 1 article).

In group (I) articles, researchers have summarized, reviewed, discussed and/or studied the (i) multiple ways of human and simian immunodeficiency viruses (HIV/SIVs) for modulation of viral/cellular gene expression, (ii) diversification (quasispecies) of Env caused by drug/immune pressure, (iii) anti-viral cellular factors against HIV/SIVs that greatly affect their biological properties, and (iv) the potential link between accessory Vpx/Vpr proteins and phylogenetic clusters of HIV/SIVs. Heusinger and Kirchhoff have described the viral/molecular mechanisms by which HIV/SIVs regulate anti-viral as well as viral gene expression in a viral replication cycle. Viral Tat protein controls the transcriptional activity of HIV/SIVs, and thus highly influences viral replication rates. Indeed, it has been demonstrated that naturally occurring Tat variations affect HIV type 1 (HIV-1) replication stage *in vivo* (Kamori and Ueno; Ronsard et al.). Also, Langer and Sauter have described various HIV-1 proteins including non-canonical ones expressed from its genome. As unique properties of HIV-1, Harada and Yoshimura, and Fujita have summarized and discussed the works on Env and on anti-viral factors in macrophages, respectively. Okada and Iwatani have described and discussed the molecular action mechanism of APOBEC3G, a major restriction factor against HIV-1. Finally, it has been reasonably postulated that adaptive mutations/variations may have contributed to the diversifications of HIV/SIVs (Sakai et al.; Sakai et al.). Miyazaki et al. has described distinct *in vitro* properties of HIV-1 and HIV-2 capsid proteins.

Needless to say, the emergence of new influenza viruses (IFVs) that are highly pathogenic to humans represents a major medical issue of great urgency. In group (II) articles, researchers have studied IFVs (mainly of the avian origin) with a special reference to their biological properties (pathogenicity for host individuals, host responses, intra- and inter-species transmission, genetic re-assortment, etc.). The variable nature of the influenza virus has been well documented, and solid anti-viral strategies based on the reported results are immediately required. The pathogenicity of H5N1, H5N6, H7N1, and H9N2 IFVs have been studied, and the results obtained have been

OPEN ACCESS

Edited by:
Aeron Hurt,
WHO Collaborating Centre for Reference and Research on Influenza (VIDRL), Australia

Reviewed by:
Ding Oh,
National University of Singapore, Singapore

****Correspondence:***
Akio Adachi
adachi@tokushima-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *19 July 2017*
Accepted: *04 September 2017*
Published: *15 September 2017*

Citation:
Nomaguchi M and Adachi A (2017) Editorial: Highly Mutable Animal RNA Viruses: Adaptation and Evolution. Front. Microbiol. 8:1785. doi: 10.3389/fmicb.2017.01785

discussed about the future direction (Jiao et al.; Gao et al.; Jin et al.; Liu et al.). While Meng et al. have studied the production of cytokines/chemokines in infected cells, Gao et al. have analyzed the host immune-related response and viral transmission efficiency in chickens. Karnbunchob et al. have focused on tracking viral avian-swine transmission, and found that most amino acid substitutions in the polymerase genes were acquired after interspecies transmission. Studies of Song et al., Yuan et al., and Zhao et al. have revealed that new viruses such as H1N2, H5N6, and H7N2, respectively, are generated by genomic reassortments. Finally, Yokoyama et al. have demonstrated, by molecular dynamics simulations, the structural basis for adaptive mutations of H3N2 IFV, a major cause for seasonal influenza in humans.

In group (III) articles, researchers have performed comprehensive genome analyses of avian or human paramyxoviruses coupled with pathogenicity studies. Satharasinghe et al. sequenced the full genome of a large number of Newcastle disease virus (NDV), showing the diversity of this virus in Malaysia. Different tissue tropisms/host ranges and pathogenicity-related host responses have been observed for various NDV strains in China (Kang et al.; Xu et al.). Saikusa et al. have systematically determined G genes of human metapneumovirus (HMPV) isolates in Japan, major causative viruses for acute respiratory diseases in humans, and have found a unique 180-nucleotide duplication in the corresponding sites of major epidemic HMPV strains. Zou et al. have determined numerous G genes of numerous respiratory syncytical virus (RSV) isolates in China, and have shown that a particular RSV genotype spreads rapidly and causes epidemics.

Flaviviruses are extraordinarily variable from biological and molecular biological points of view and are the topic of group (IV) articles. Hu et al. have determined E2 genes derived from numerous classical swine fever virus (CSFV) isolates from C-strain-vaccinated pigs to know the etiological reason for sporadic CSF outbreaks in the area. The results have shown that circulating CSFVs in those pigs mostly belong to a specific sub-genotype. Kim et al. have examined a retinoid-interferon-induced cell-mortality factor designated GRIM-19 for its effect on hepatitis C virus (HCV) replication, and suggested that GRIM-19 acts as an anti-viral host factor by reducing intracellular lipid accumulation. A virus designated Zika (ZIKV) has recently become the focus of public attention due to a remarkable number of microcephaly cases in Brazil. As for ZIKV, Saiz et al. have extensively reviewed and summarized the present knowledge, from basic information on biology and molecular biology, medical issues, and through to public health matters. Koide et al. have reported that the cynomolgus macaque can serve as an infection model for ZIKV.

In group (V) to (VIII) articles, rabies virus (RABV), noroviruses of murine (MuNoV) or human (NoV) origin, picornaviruses (enterovirus species C (EV-C) and human rhinovirus C (HRV-C)), and picobirnaviruses (PBVs) have been studied. Jamalkandi et al. have started the systems biology to identify a network of proteins significant in the RABV infection. The importance of RABV P protein for viral replication and pathogenicity has been shown by genetic manipulation as expected (Mei et al.). Kitamoto et al. have focused on the laboratory adaptation of MuNoV, and have suggested that the interplay between variants is necessary for the virus to better adapt for growth in cell culture. Sato et al. have performed an integrated analysis on NoV GII.4, a major cause of viral gastroenteritis in humans. They have examined the pandemic lineage for its molecular evolution by using viral full-length genome and VP1 sequences from a large number of samples collected in Japan. The results are important to understand the biology of NoV and to control NoV infection. Bessaud et al. have successfully developed a method to sequence rapidly the entire EV-C genome. It is useful to identify the type of EV-C strain and also to distinguish it from closely related viruses. Khaw et al. have determined the complete genome sequence of seven Malaysian HRV-C isolates, and found unique and conserved sequences relative to that of HRV-Cs from the other countries. Finally, Woo et al. have determined partial sequences of PBVs isolated from numerous animal species, and performed a phylogenetic analysis. PBVs have been recently discovered in a wide variety of mammals and birds. The results obtained have demonstrated a highly divergent nature of PBV.

As mentioned in the first part of this article, there are 36 papers in this Research Topic covering a wide variety of animal RNA viruses. Each work has its own scientific impact, and highlights virological significance and interest of the virus diversification. Animal RNA viruses readily mutate in response to certain biological and/or chemical stimuli rendered by the surrounding environments. To precisely know or understand the biological processes and the underlying molecular mechanisms remains a mission for virologists, requiring both applied as well as basic virology.

AUTHOR CONTRIBUTIONS

MN and AA wrote the manuscript, and approved its submission.

FUNDING

This work is supported by Research Program on HIV/AIDS Grant Numbers 15545611 and 16768720 to AA and MN, respectively, from Japan Agency for Medical Research and Development (AMED).

ACKNOWLEDGMENTS

We thank all the contributors to this Research Topic for their intriguing works. We also thank Ms. Kazuko Yoshida (Tokushima University Graduate School of Medical Science) for editorial assistance.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Nomaguchi and Adachi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

REVIEW
published: 14 February 2017
doi: 10.3389/fmicb.2017.00198

Primate Lentiviruses Modulate NF-κB Activity by Multiple Mechanisms to Fine-Tune Viral and Cellular Gene Expression

*Elena Heusinger and Frank Kirchhoff**

Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany

The transcription factor nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) plays a complex role during the replication of primate lentiviruses. On the one hand, NF-κB is essential for induction of efficient proviral gene expression. On the other hand, this transcription factor contributes to the innate immune response and induces expression of numerous cellular antiviral genes. Recent data suggest that primate lentiviruses cope with this challenge by boosting NF-κB activity early during the replication cycle to initiate Tat-driven viral transcription and suppressing it at later stages to minimize antiviral gene expression. Human and simian immunodeficiency viruses (HIV and SIV, respectively) initially exploit their accessory Nef protein to increase the responsiveness of infected $CD4^+$ T cells to stimulation. Increased NF-κB activity initiates Tat expression and productive replication. These events happen quickly after infection since Nef is rapidly expressed at high levels. Later during infection, Nef proteins of HIV-2 and most SIVs exert a very different effect: by down-modulating the CD3 receptor, an essential factor for T cell receptor (TCR) signaling, they prevent stimulation of $CD4^+$ T cells via antigen-presenting cells and hence suppress further induction of NF-κB and an effective antiviral immune response. Efficient LTR-driven viral transcription is maintained because it is largely independent of NF-κB in the presence of Tat. In contrast, human immunodeficiency virus type 1 (HIV-1) and its simian precursors have lost the CD3 down-modulation function of Nef and use the late viral protein U (Vpu) to inhibit NF-κB activity by suppressing its nuclear translocation. In this review, we discuss how HIV-1 and other primate lentiviruses might balance viral and antiviral gene expression through a tight temporal regulation of NF-κB activity throughout their replication cycle.

Keywords: HIV, SIV, NF-κB, Nef, Vpu, Tat, LTR

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Mikako Fujita,
Kumamoto University, Japan
Takamasa Ueno,
Kumamoto University, Japan
Kenzo Tokunaga,
National Institute of Infectious Diseases, Japan

****Correspondence:***
Frank Kirchhoff
frank.kirchhoff@uni-ulm.de

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *04 January 2017*
Accepted: *27 January 2017*
Published: *14 February 2017*

Citation:
Heusinger E and Kirchhoff F (2017) Primate Lentiviruses Modulate NF-κB Activity by Multiple Mechanisms to Fine-Tune Viral and Cellular Gene Expression. Front. Microbiol. 8:198. doi: 10.3389/fmicb.2017.00198

INTRODUCTION

To allow efficient viral gene expression, replication, and spread, viral pathogens need to exploit the cellular transcriptional machinery. In some cases, they hijack exactly those transcription factors that are activated by the host response to infection to initiate antiviral immune responses. The interaction of human immunodeficiency virus type 1 (HIV-1) and related simian immunodeficiency viruses (SIVs) with the NF-κB (nuclear factor kappa-light-chain-enhancer of

activated B cells) family of transcription factors provides a prime example for the hostile takeover of a key mediator of the immune response by viral pathogens (Hiscott et al., 2001; Chan and Greene, 2012). NF-κB is ubiquitously expressed and its dysregulation is associated with many pathologies including cancer, cardiovascular, pulmonary, and inflammatory diseases (Panday et al., 2016). NF-κB can be induced by multiple stimuli and regulates the expression of cellular genes involved in numerous processes, such as cell proliferation, DNA repair and cell differentiation or survival (Ghosh and Hayden, 2012; Napetschnig and Wu, 2013). Furthermore, NF-κB plays a key role in inflammation and the induction of innate and adaptive immune responses, including expression of interferon-stimulated genes (ISGs) representing a first line of defense against viral pathogens (Hayden et al., 2006; Vallabhapurapu and Karin, 2009; Pfeffer, 2011).

NF-κB has a complex role in HIV-1 replication and pathogenesis. ISGs exert numerous effector functions and may target almost every step of the retroviral replication cycle (Kluge et al., 2015; Colomer-Lluch et al., 2016). Thus, their induction by NF-κB transcription factors might have beneficial effects for the host by suppressing HIV-1 replication. However, induction of interferon and other cytokines also contributes to HIV-1-induced chronic and systemic inflammation that drives progression to AIDS. Thus, whether induction of interferon responses is beneficial or harmful in HIV-1 infection is a matter of debate (Doyle et al., 2015; Utay and Douek, 2016). Importantly, NF-κB is also critical for potent viral gene expression from the HIV-1 long terminal repeat (LTR) promoter that typically contains two adjacent NF-κB binding sites in its main enhancer region (Chan and Greene, 2012). Combinations of stimulatory drugs including NF-κB inducing agents, such as prostratin, are currently examined for their ability to activate the latent reservoirs of HIV-1 to render them susceptible for elimination (Jiang and Dandekar, 2015; Cary et al., 2016). Thus, NF-κB presents a target for therapeutic intervention both by suppressing its activation to prevent harmful inflammation (Chan and Greene, 2012) or by activating this transcription factor to stimulate the latent reservoirs of HIV-1 (Jiang and Dandekar, 2015).

Numerous studies have examined the effects of HIV-1 infection on NF-κB activity (reviewed in Chan and Greene, 2012). Altogether, the results were puzzling and frequently controversial. Recent evidence suggests that HIV-1 and other primate lentiviruses may have differential effects during the early and late stages of their replication cycle and tightly regulate NF-κB activity (Sauter et al., 2015). In the present review, we first briefly describe some basic aspects of NF-κB signaling and its interaction with the LTR promoter of HIV-1 and other primate lentiviruses. Subsequently, we discuss how these viruses might modulate NF-κB activity throughout their replication cycle to ensure efficient viral gene expression while minimizing the expression of antiviral genes.

BASIC MECHANISMS OF NF-κB SIGNALING

NF-κB was initially discovered in the laboratory of David Baltimore because of its ability to bind the enhancer region of the immunoglobulin κ light-chain- in B cells (Sen and Baltimore, 1986). Since then, five mammalian NF-κB members, i.e., NF-κB1 (p50/p105), NF-κB2 (p52/p100), RelA (p65), RelB and c-Rel, have been identified (Nabel and Verma, 1993). All of them share a highly conserved N-terminal Rel homology domain that is critical for dimerization and DNA binding, while three of them (p65, RelB, and c-Rel) contain an additional C-terminal transactivation domain. The remaining two factors (NF-κB1 and NF-κB2) are synthesized as large inactive precursors (p105 and p100) that need to undergo proteasomal degradation of their ankyrin repeat containing C-terminal region to generate the mature p50 and p52 NF-κB subunits, respectively (Karin and Ben-Neriah, 2000). DNA binding and transcriptional regulation requires dimerization of two subunits, with p50/p65 dimers being the most abundant active form. Many other dimer combinations were observed but not all of them act as transcriptional activators; e.g., p50/p50 and p52/p52 homodimers were reported to suppress NF-κB mediated transcription as they lack a transactivation domain (Zhong et al., 2002).

In unstimulated cells, NF-κB dimers are bound to so-called inhibitors of κB (IκBs) containing ankyrin repeats that mask nuclear localization signals (NLS) and thereby keep the NF-κB proteins sequestered in the cytoplasm. The p105 and p100 precursors also contain ankyrin repeats in their C-terminal parts. Thus, they inhibit NF-κB activity and may also be classified as IκB proteins. A multitude of stimuli including viral antigens, immune cells, mitogens and cytokines, specific cell surface receptors, such as the T cell receptor (TCR)-CD3 complex, IFN receptors, the interleukin 1 receptor (IL-1R), Toll-like receptors (TLRs), or tumor necrosis factor (TNF) receptors induce degradation of IκB proteins (Vallabhapurapu and Karin, 2009). The main mechanism (canonical pathway) involves activation of IκB kinase (IKK), a heterodimer composed of the catalytic IKKα and IKKβ subunits and a regulatory factor termed NEMO (NF-κB essential modulator), and subsequent phosphorylation of serine residues in IκB regulatory domains allowing its ubiquitination and proteasomal degradation (**Figure 1**). Degradation of IκB allows exposure of the NLS and thus entry of the NF-κB complex into the nucleus for DNA binding and transcriptional induction of genes containing the appropriate binding sites within their promoter regions. As outlined below, NF-κB activation induces immune responses as well as HIV-1 proviral gene expression. Notably, NF-κB also induces expression of IκBα, resulting in a negative autoregulatory feedback loop and thus oscillating levels of NF-κB activity (Nelson et al., 2004). More in depth reviews of NF-κB signaling and its regulation are provided by Oeckinghaus and Ghosh (2009), Skaug et al. (2009) and Vallabhapurapu and Karin (2009).

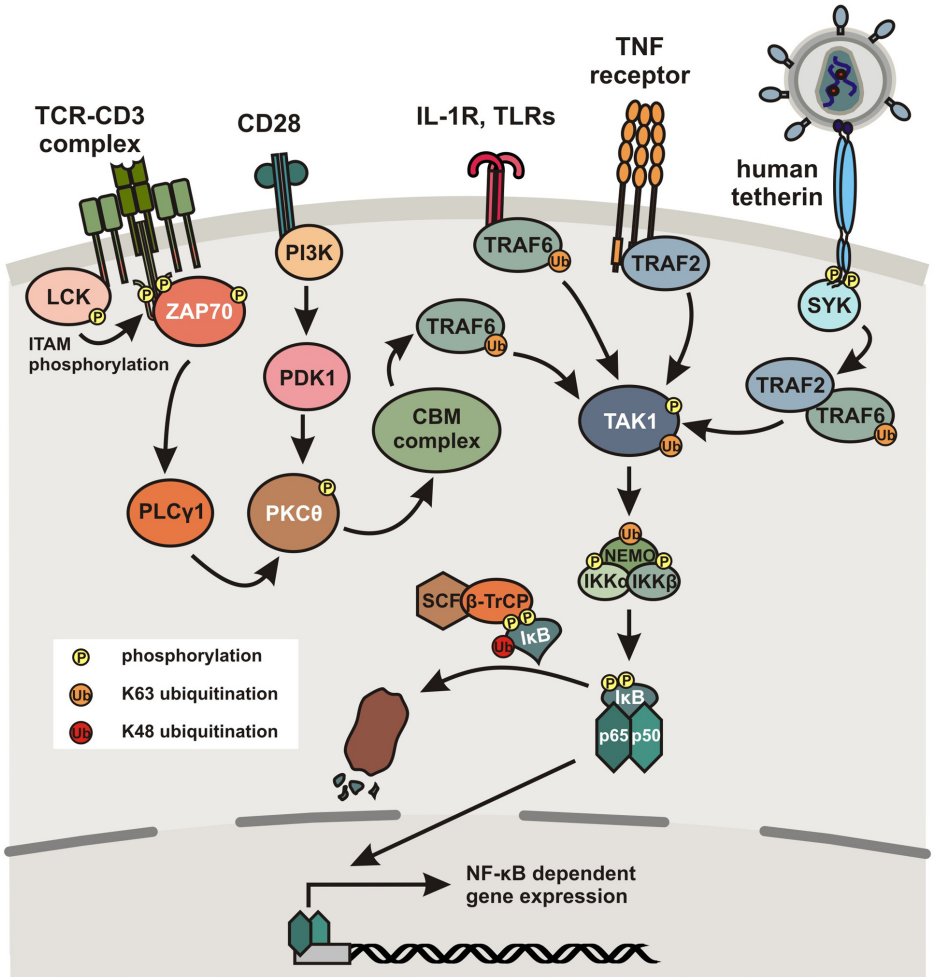


FIGURE 1 | Activation of the canonical NF-κB pathway. Simplified schematic presentation of NF-κB activation by different cellular receptors. Binding of the TCR-CD3 complex to antigens presented by MHC molecules induces LCK-dependent phosphorylation of ITAMs in the cytoplasmic tail of CD3. This in turn recruits ZAP70, which induces a signal cascade involving activation of PLCγ1 and PKCθ. Alternatively, PKCθ can also be activated via a signal cascade including PI3K following binding of the TCR costimulatory factor CD28 to a B7 ligand. PKCθ subsequently activates the CBM complex consisting of CARD11, BCL10 and MALT1, which activates TAK1 in a TRAF6 dependent manner. Ligand binding to TLRs or TNF receptors also induces TAK1 activation in a TRAF6- or TRAF2-dependent manner, respectively. In humans (and to a much lesser extent in chimpanzees), the cellular restriction factor tetherin also acquired the ability to activate NF-κB. Upon tethering of newly formed viral particles and clustering of tetherin dimers, the cytoplasmic tail of tetherin becomes phosphorylated at a conserved YxY motif by a Scr-family kinase in a RICH2-dependent manner, recruiting SYK, which initiates a signaling cascade involving TRAF2 and/or TRAF6 resulting in the activation of TAK1. TAK1 subsequently phosphorylates IKKβ at two serine residues in the activation loop to activate IKK. Furthermore, TRAFs activate NEMO by poly-ubiquitination, followed by activation of the catalytic subunits IKKα and IKKβ. Subsequently, IκB is phosphorylated and poly-ubiquitinated by the SCF/β-TrCP complex, which results in its proteasomal degradation, thereby releasing p50/p65 heterodimers, which translocate to the nucleus where they bind to specific κB binding sites and initiate NF-κB-dependent gene expression.

ROLE OF NF-κB IN IMMUNITY AND INFLAMMATION

It is well established that NF-κB transcription factors regulate many genes involved in immune and inflammatory responses (reviewed in Vallabhapurapu and Karin, 2009). Upon pathogen recognition, a variety of cellular pattern recognition receptors (PRRs), including cGAS [Cyclic-GMP-AMP (cGAMP) synthase], STING (stimulator of interferon genes), IFI16 (interferon γ-inducible protein 16), RIG-I-like helicases, NOD-like receptors (NLRs) and TLRs may activate the NF-κB pathway. Thus, NF-κB signaling is induced by multiple stimuli, including pathogen associated molecular patterns (PAMPs) as well as proinflammatory cytokines, e.g., tumor necrosis factor-α (TNF-α) and interleukin-1. Ligand binding to TLRs induces several downstream effectors and results in the formation of signaling complexes activating the IKK complex (**Figure 1**). Ubiquitination of TNF-receptor-associated factors (TRAF) and NEMO facilitates activation of the catalytic IKKβ subunit resulting in the phosphorylation and proteasomal degradation of IκB and consequently nuclear translocation of active NF-κB dimers (Skaug et al., 2009). Similarly, signaling via members of the

TNF receptor superfamily ultimately leads to activation of the IKK complex, following recruitment of various adaptor proteins and stimulation of TGF-β-activated kinase 1 (TAK1; Wang and Baldwin, 1998). Furthermore, some activated TNFR superfamily members may induce accumulation of mitogen-activated protein kinase 14 (MAP3K14 or NF-κB-inducing kinase, NIK) that activates IKKα to induce processing of p100 and thus activation of e.g., RelB/p52 heterodimers to mediate expression of NF-κB-responsive genes via the non-canonical pathway. This pathway is slower than the canonical NF-κB signaling pathway and can also be activated via the CD40 ligand (CD40L), receptor activator of NF-κB ligand (RANKL) or lymphotoxin β (Cildir et al., 2016). Following activation, RelB/p52 heterodimers translocate to the nucleus for target gene activation.

NF-κB is not only a key factor in the induction of effective innate immune responses but also plays an important role in adaptive immunity. For example, ligand binding to the TCR complex induces recruitment of LCK tyrosine kinase that phosphorylates the ITAMs of the CD3 ζ chains (**Figure 1**). Subsequent recruitment of ZAP70 mediates activation of a signaling pathway involving PLCγ, protein kinase C (PKC) family members, the CARD11-BCL10-MALT1 (CBM) complex, TRAF6, and ultimately TAK1, inducing NF-κB activity by phosphorylation and activation of IKKs (Cheng et al., 2011; Paul and Schaefer, 2013). NF-κB activation is further enhanced by costimulation via CD28, which triggers distinct signaling cascades involving PI3K and PDK1 (Boomer and Green, 2010). Thus, the signaling pathways induced by CD3/CD28-mediated stimulation of T cells upon interaction with antigen-presenting cells (APCs) allow NF-κB to enter the nucleus to upregulate expression of cytokines and antimicrobial effectors as well as genes involved in T cell survival, proliferation, and differentiation. Consequently, regulation of NF-κB activity plays a key role in innate and adaptive immune function and the defense against bacterial and viral pathogens. In agreement with this important role in inflammatory gene expression, chronic activation of NF-κB signaling is observed in a variety of inflammatory diseases including arthritis, sepsis, gastritis, asthma, atherosclerosis, and inflammatory bowel disease (Panday et al., 2016).

ROLE OF NF-κB IN PRIMATE LENTIVIRAL TRANSCRIPTION

It is long known that HIV-1 transcription can be stimulated by activation of the canonical NF-κB pathway (Nabel and Baltimore, 1987; Tong-Starksen et al., 1987). NF-κB binding sites are found in the enhancer region of all primate lentiviral LTRs, although their numbers may vary between different subtypes of HIV-1 group M and various groups of SIV and HIV. Most subtypes of pandemic HIV-1 group M strains (A, B, D, F, G, H, J, and K) and some SIVs contain two NF-κB binding sites located –104 to –80 bp upstream of the transcriptional start site within the 5′LTR (**Figure 2**). In contrast, the second human immunodeficiency virus (HIV-2), subtype A/E recombinants of HIV-1 group M and several SIV lineages contain just a single NF-κB binding site. Finally, subtype C strains, which account for almost 50%

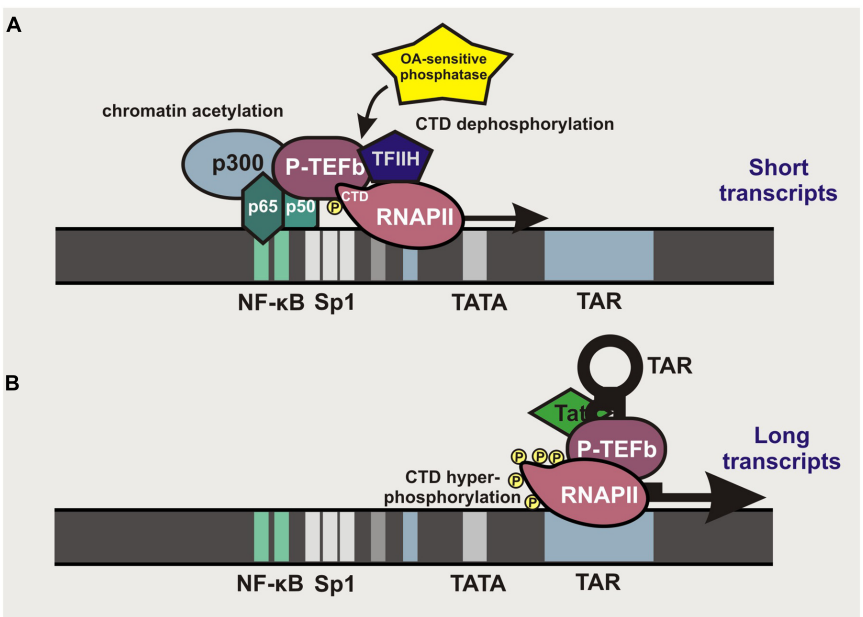


FIGURE 2 | Role of κB binding sites in primate lentiviral LTR-mediated transcription. (A) Localization of κB binding sites in primate lentiviral LTRs. NF-κB binding allows recruitment of p300 to initiate chromatin acetylation and to render the LTR better accessible for RNAPII. NF-κB also recruits P-TEFb, which binds to the CTD of RNAPII and strongly enhances its processivity. De-phosphorylation of the CTD by an OA-sensitive phosphatase terminates the elongation, resulting in the production of short TAR-containing transcripts. **(B)** In the presence of the viral Tat protein, transcription of proviral DNA is maintained independently of NF-κB. Tat binds to the short hairpin loop of TAR and recruits P-TEFb to the RNAPII, thereby allowing efficient elongation and generation of full-length viral transcripts.

of HIV-1 infections worldwide, typically contain three binding sites for NF-κB in their enhancer region (Bachu et al., 2012). LTR-mediated transcription is initiated by binding of p50/p65 heterodimers or other members of the NF-κB transcription factor family that recruit p300 to promote acetylation of LTR chromatin and allow access for RNA polymerase II (RNAPII; **Figure 2A**). NF-κB binding also promotes transcriptional elongation by recruiting the positive transcription elongation factor b (P-TEFb) complex and the transcription factor TFIIH to the carboxyl domain (CTD) of RNAPII, allowing its phosphorylation that results in increased processivity of this enzyme (Barboric et al., 2001). However, this effect is transient since an okadaic acid (OA)-sensitive phosphatase dephosphorylates the CTD of RNAPII to prevent P-TEFb interaction. Thus, short transcripts terminated shortly after the trans-activation response (TAR) RNA element predominate at the early stage. Accumulation of the viral Tat protein that binds to TAR and recruits P-TEFb allows the virus to overcome this bottleneck by allowing hyper-phosphorylation of RNAPII and potent transcriptional elongation and the generation of full-length viral transcripts (Williams et al., 2007) (**Figure 2B**).

Typically, mutation of the NF-κB binding sites in HIV-1 LTRs will prevent efficient proviral transcription. However, one case of a replication-competent pathogenic HIV-1 strain lacking NF-κB binding sites has been reported (Zhang et al., 1997). This unusual HIV-1 strain contained duplications in the TCF-1alpha region that may have compensatory effects. Thus, NF-κB binding sites are important but not essential for HIV-1 replication or pathogenicity *in vivo*. Since NF-κB activation stimulates HIV-1 transcription, it is also targeted in approaches aiming to eliminate the latent viral reservoirs. Induction of NF-κB activity by T cell activation or treatment with phorbol esters, such as prostratin, potently enhance viral gene expression (Coudronniere et al., 2000; Williams et al., 2004). However, since NF-κB transcription factors are involved in numerous physiological and pathological processes, their induction is prone to adverse effects. Notably, some members of the NF-κB transcription factor family may also promote HIV-1 latency, i.e., it has been reported that in unstimulated T cells, p50 homodimers occupy the NF-κB sites in the proviral LTR to recruit HDAC1, thereby promoting histone hypo-acetylation and hence chromatin condensation rendering the viral LTR poorly accessible to RNAPII binding (Williams et al., 2006).

INDUCTION OF NF-κB SIGNALING BY VIRAL IMMUNE SENSING AND DNA DAMAGE RESPONSES

HIV-1 infection modulates NF-κB signaling by multiple mechanisms. As outlined above, sensing of viral PAMPs by PRRs activates NF-κB to induce an antiviral immune response. HIV-1 sensing is not fully understood as the virus has evolved effective evasion mechanisms. For example, recent evidence suggests that the viral capsid stays largely intact until it docks to the nucleopore (Arhel et al., 2007; Lelek et al., 2012; Dharan et al., 2016) and that it recruits cellular factors, such as PSF6 and cyclophilins, to prevent innate immune activation (Rasaiyaah et al., 2013). Nonetheless, it has become clear that a variety of viral components may trigger immune responses. Viral RNA or DNA intermediates of the reverse transcription (RT) process can be sensed by cytosolic receptors, such as cGAS, IFI16, PQBP1 and RIG-I, particularly under suboptimal conditions for efficient RT (Jakobsen et al., 2015; Sauter and Kirchhoff, 2016). Moreover, antiretroviral restriction factors may also act as immune sensors (Hotter et al., 2013). For example, TRIM5α induces untimely uncoating of the viral capsid and may also act as an activator of the TAK1 kinase complex to stimulate AP-1 and NF-κB signaling (Pertel et al., 2011). Similarly, trapping of HIV-1 particles by the host restriction factor tetherin induces phosphorylation of tyrosine residues in its cytoplasmic tail to mediate recruitment of SYK tyrosine kinase and TRAF2 and/or 6 to activate TAK1 and consequently NF-κB-dependent immune responses (Galão et al., 2012, 2014; Tokarev et al., 2013) (**Figure 1**). Interestingly, this sensing function of tetherin seems to have an evolutionary recent origin and is only observed for the human and (to a much lesser extent) chimpanzee orthologs (Galão et al., 2012), whereas the ability of tetherin to block virion release has a very ancient origin (Heusinger et al., 2015; Blanco-Melo et al., 2016). Finally, HIV-1 infection induces DNA damage responses since the generation of linear viral RNA/DNA and DNA species, non-integrated circular forms of viral DNA and a double-stand break in the host-genome are inevitable concomitants of the viral replication cycle. DNA damage signaling may activate ATM and mediate phosphorylation and ubiquitination of NEMO to induce IKK and consequently NF-κB activation (Miyamoto, 2011; McCool and Miyamoto, 2012).

ACTIVATION OF NF-κB TO INITIATE EARLY VIRAL GENE TRANSCRIPTION

HIV-1 not only affects NF-κB activation by triggering immune sensors and inducing DNA damage responses but may also use some of its gene products to manipulate this transcription factor to promote efficient viral gene expression. Many studies investigated the effect of HIV-1 on NF-κB activation but the initial results were often contradictory. One possible reason for this is that NF-κB signaling plays a complex role in the viral replication cycle and that effects may depend on the cell type and state of activation, as well as the stimuli and HIV-1 strains or proteins used in the respective studies. Furthermore, accumulating evidence suggests that the HIV-1 Nef and Tat proteins that are expressed at high levels immediately after initiation of proviral transcription enhance and late viral gene products, such as Vpu, suppress NF-κB activation (Roux et al., 2000; Akari et al., 2001; Bour et al., 2001; Varin et al., 2005; Herbein et al., 2008; Mangino et al., 2011; Fiume et al., 2012; Liu et al., 2013). It has been reported that Tat interacts with IκBα and the p65 subunit of NF-κB to prevent binding of the repressor to the NF-κB complex while promoting p65 binding to DNA (Fiume et al., 2012). It has also been shown that the cytoplasmic domain of the HIV-1 envelope glycoprotein (Env) gp41 interacts with TAK1 to induce NF-κB activation (Postler and Desrosiers,

2012). Whether this mechanism is effective during the earliest stage of infection, i.e., induced by virion fusion with the plasma cell membrane of the target cell, or requires higher quantities of Env achieved during later stages of the replication cycle remains to be determined. Finally, a variety of studies reported that Vpr modulates NF-κB signaling in various cell types. However, the effects are controversial and both stimulatory and inhibitory effects of Vpr have been described (Kogan et al., 2013; Liu et al., 2013, 2014; Liang et al., 2015). Thus, the effects of virion-associated and *de novo* synthesized Vpr in primary HIV-1 infected T cells need to be further investigated.

The accessory viral Nef protein does not induce NF-κB activity on its own but boosts the responsiveness of HIV-1 infected cells to stimulation (Alexander et al., 1997; Simmons et al., 2001; Fenard et al., 2005) (**Figure 3A**). Nef-mediated activation of NF-κB, nuclear factor of activated T cells (NFAT), IL-2 and LTR stimulation following TCR-CD3/CD28 costimulation seems to require association with lipid rafts (Schrager and Marsh, 1999; Fortin et al., 2004; Kumar et al., 2016) and may depend on the state of activation of infected T cells (Baur et al., 1994). Interactions of HIV-1 Nef with the CD3 ζ chain (Xu et al., 1999) and downstream effectors of TCR signaling, such as the tyrosine kinase LCK (Baur et al., 1997), serine/threonine p21-activating kinases (Sawai et al., 1994), the DOCK2-ELMO1 complex (Janardhan et al., 2004) and ERK/MAP kinases (Schrager et al., 2002) have been reported. Thus, Nef might affect the catalytic activity of different kinases, induce cytoskeletal changes, and activate a variety of signaling pathways. The relative contribution of these activities and interactions to Nef-mediated enhancement of T cell activation is largely unclear. In either case, Nef promotes nuclear translocation of NF-κB and other transcription factors, such as AP1 and NFAT and activates the viral promoter to induce Tat expression and hence productive viral replication (Kinoshita et al., 1998). Notably, Nef may exert its multiple functions rapidly after viral entry since it is expressed at high levels early during the viral replication cycle and possibly even before proviral integration (Sloan et al., 2011).

LATE INHIBITION OF NF-κB SIGNALING BY HIV-1 AND ITS *VPU* CONTAINING SIV COUNTERPARTS

While Nef-mediated enhancement of NF-κB activity may be important to initiate proviral transcription, it is less critical after accumulation of the viral Tat protein and may even become detrimental to HIV-1 replication because of the induction of antiviral gene expression. Thus, HIV-1 and other primate lentiviruses might tightly regulate NF-κB activity throughout their replication cycle to allow proviral transcription while minimizing antiviral gene expression. Recent studies show that the accessory viral protein U (Vpu) potently suppresses NF-κB activity during later stages of the viral replication cycle (**Figure 3B**) (Akari et al., 2001; Bour et al., 2001; Sauter et al., 2015). A *vpu* gene was most likely acquired by the precursor of SIVs infecting *Cercopithecus* monkeys with subsequent cross-species transmissions and recombination events giving rise to other *vpu* containing primate lentiviruses (Bailes et al., 2003; Takeuchi et al., 2015). Thus, *vpu* is found in HIV-1, its chimpanzee and gorilla precursors, SIVcpz and SIVgor, and in SIVgsn, SIVmus, SIVmon, and SIVden, infecting greater spot-nosed, mustached, mona, and Dent's mona monkeys, respectively (Kirchhoff, 2009). The Vpu proteins of pandemic HIV-1 group M strains counteract tetherin-mediated activation of NF-κB-dependent antiviral immune responses (Cocka and Bates, 2012; Galão et al., 2012, 2014; Tokarev et al., 2013). This is expected since the Vpus of HIV-1 group M and SIVs infecting several *Cercopithecus* species are potent antagonists of tetherin-mediated inhibition of virion release. More surprisingly, Vpus derived from SIVs infecting chimpanzees and gorillas or HIV-1 group O strains that use their Nef protein to counteract tetherin are also potent inhibitors of NF-κB (Sauter et al., 2015). Indeed, early studies suggested that HIV-1 Vpu prevents NF-κB activation by inhibiting degradation of IκB through sequestration of the adaptor protein β-TrCP (Akari et al., 2001; Bour et al., 2001). More recent data show that the ability of Vpu to prevent NF-κB activation independently of the stimulus is conserved between all lineages of SIV and HIV-1 (except group N) containing this accessory gene and does not correlate with β-TrCP interaction or tetherin antagonism (Sauter et al., 2015). In agreement with the results of early studies (Bour et al., 2001), an intact β-TrCP binding motif and interaction of Vpu with β-TrCP are essential for highly effective inhibition of p65 nuclear translocation by primary Vpu proteins. However, some Vpus failing to recruit β-TrCP still suppressed NF-κB-dependent gene expression (Sauter et al., 2015). Thus, the action of Vpu involves stabilization of IκB and reduced nuclear translocation of p65 but also additional yet-to-be-defined mechanisms. Importantly, the ability of Vpu to inhibit NF-κB is dominant over the stimulatory effect of Nef and associated with reduced induction of IFN and ISGs in HIV-1 infected T cells (Sauter et al., 2015). Accordingly, HIV-1 and its closest simian counterparts seem to utilize Nef to boost NF-κB activation to initiate LTR-driven proviral transcription. At later stages, when viral gene expression is ensured by the presence of Tat, Vpu inhibits NF-κB activity to limit expression of antiviral genes and to attenuate the immune response.

POTENTIAL SUPPRESSION OF NF-κB ACTIVATION BY NEF-MEDIATED DOWN-MODULATION OF TCR-CD3

The ability to inhibit NF-κB activity is highly conserved among primate lentiviral Vpu proteins (Sauter et al., 2015) suggesting an important role for viral immune evasion *in vivo*. As outlined above, however, *vpu* genes are only found in HIV-1 and its closest simian counterparts, raising the question whether the majority of primate lentiviruses use another mechanism to inhibit NF-κB during late stages of the replication cycle. It is tempting to speculate that originally essentially all primate lentiviruses

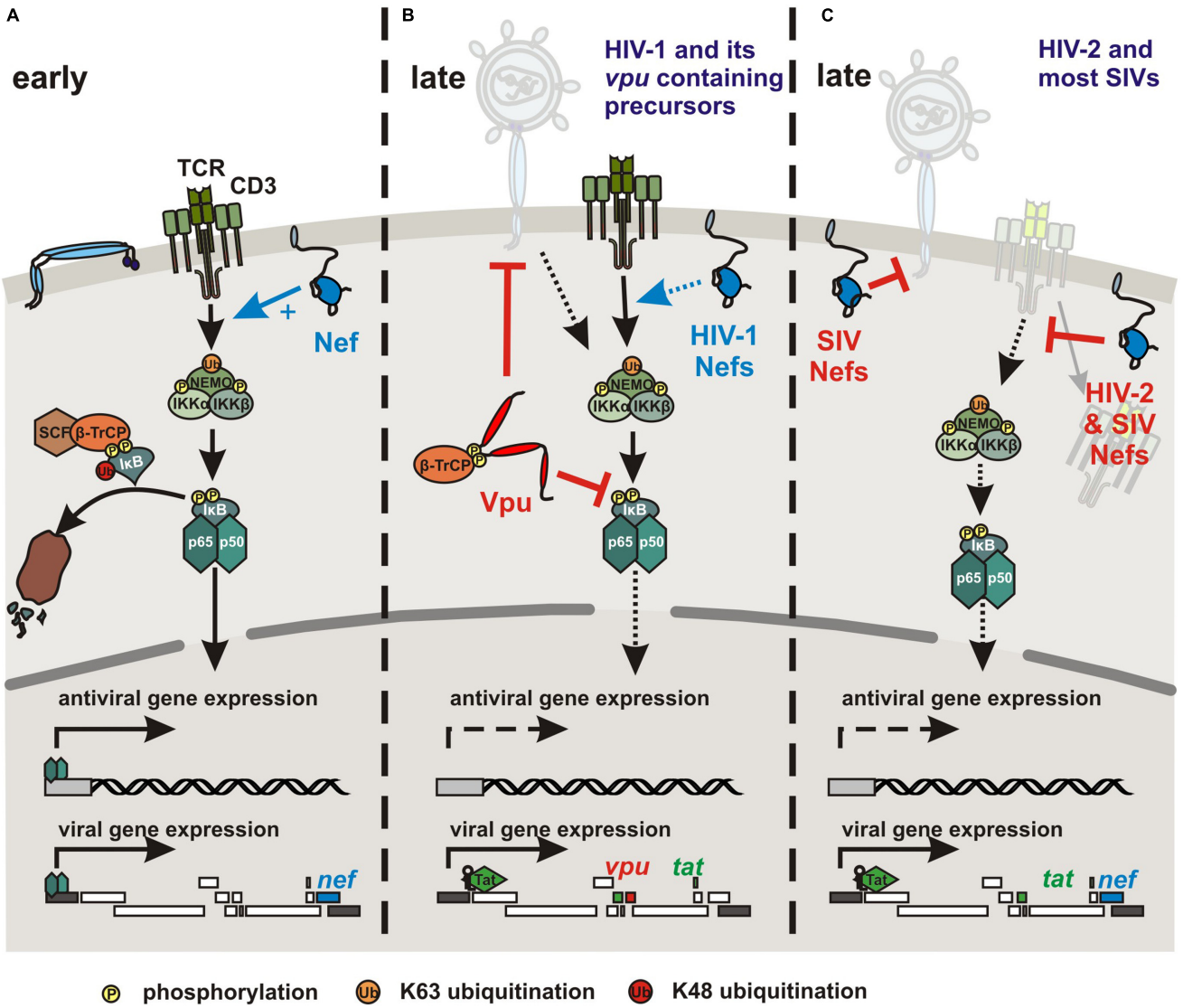


FIGURE 3 | Modulation of NF-κB activity by HIV and SIV Nef and Vpu proteins. (A) To initiate early viral gene expression, the viral accessory protein Nef promotes NF-κB activation by boosting TCR-CD3 mediated T cell activation and other yet to be determined mechanisms. **(B)** HIV-1 and its closest SIV counterparts use their Vpu protein to inhibit NF-κB and thus antiviral gene expression during late stages of the replication cycle. Vpu interferes with IκB degradation by sequestration of β-TrCP and other as-yet-unknown mechanisms. Furthermore, HIV-1 group M and (less effectively) N Vpu proteins counteract the cellular restriction factor tetherin, which traps budding virions at the cell surface and also acts as NF-κB activating immune sensor in the case of the human ortholog. In the presence of the viral transactivator Tat, viral transcription is maintained independently of NF-κB activity. **(C)** HIV-2 and most SIV strains do not contain a *vpu* gene but express Nef proteins that efficiently down-modulate CD3 from the cell surface to prevent T cell activation and hence to suppress the induction of NF-κB and other transcription factors. SIV Nefs also counteract tetherin in their respective host. However, although monkey tetherins restrict virus release they are not known to act as NF-κB activating immune sensors.

used the CD3 down-modulation function of Nef to suppress the activity of NF-κB and other transcription factors. This Nef function is conserved among HIV-2 and most lineages of SIVs but was entirely lost in the great majority of *vpu* containing primate lentiviruses (Schindler et al., 2006). This was most likely not just coincidence. In fact, phylogenetic and functional analyses suggest that the CD3 down-modulation function of Nef may have been lost twice during primate lentiviral evolution when the virus acquired a *vpu* gene. The first time after acquisition of *vpu* by a precursor of SIVs nowadays found in various *Cercopithecus* monkeys and a second time when this virus recombined with the precursor of SIVrcm from Red-capped mangabeys in chimpanzees to give rise to SIVcpz, the precursor of HIV-1 (Kirchhoff, 2010).

It has been suggested that Vpu might diminish the selective advantage of Nef-mediated CD3 down-modulation because it counteracts tetherin and potentially other antiviral factors induced in an inflammatory environment. In fact, lack of CD3

down-modulation by Nef is associated with higher expression levels of early (CD69) and late (CD25) activation markers as well as increased levels of apoptosis, induction of Fas, Fas-L, PD-1, and CTLA-4 cell surface expression and secretion of interferon gamma (IFN-γ) in virally infected cultures of peripheral blood mononuclear cells (Schindler et al., 2006; Schmökel et al., 2011; Yu et al., 2015). The consequences of the presence of *vpu* and the inability of Nef to block TCR-mediated T cell activation, that distinguish HIV-1 and its simian precursors from other primate lentiviruses, for viral pathogenicity largely remain a matter of speculation. While it is evident that host properties play an important role in the clinical outcome of primate lentiviral infections (Chahroudi et al., 2012), it is conceivable that potent inhibition of T cell activation should help to prevent damagingly high levels of immune activation that drive $CD4^+$ T cell depletion and progression to immunodeficiency (Sodora et al., 2009). Indeed, inefficient Nef-mediated down-modulation of CD3 correlates with low numbers of $CD4^+$ T cells in SIVsmm infected sooty mangabeys (Schindler et al., 2008) and viremic HIV-2 infected individuals (Khalid et al., 2012). Furthermore, HIV-1 is more pathogenic than HIV-2 (Nyamweya et al., 2013) and many SIVs do not cause disease in their natural host species (Chahroudi et al., 2012), whereas SIVcpz causes AIDS in chimpanzees (Keele et al., 2009). It is unknown whether other *vpu* containing viruses cause disease in the wild. However, the prevalence of SIVgsn/mus/mon in their natural simian hosts seems to be much lower (∼1–4%) than of SIVs capable of CD3 down-modulation (often >40%; Heigele et al., 2016). It will be interesting to further examine whether this is due to differences in the pathogenic outcome of these infections.

Vpu may allow primate lentiviruses to better cope with the antiviral immune response not only by antagonizing innate antiviral factors but also by downregulation of various receptors involved in the activation of natural killer cells (Sugden et al., 2016). However, the recent finding that Vpu suppresses NF-κB activity also suggests a more direct link with the CD3 down-modulation function of Nef. As described above, signaling via the TCR-CD3 complex induces a cascade of events ultimately leading to IKK activation and translocation of NF-κB into the nucleus for target gene expression (**Figure 1**). It has been established that Nef-mediated down-modulation of CD3 potently blocks the responsiveness of virally infected T cells to TCR-mediated stimulation (Iafrate et al., 1997; Bell et al., 1998; Khalid et al., 2012) and prevents the formation of the immunological synapse between virally infected primary $CD4^+$ T cells and dendritic cells or macrophages (Arhel et al., 2009). It has been shown that CD3 down-modulation potently inhibits the induction of NFAT (Khalid et al., 2012), which also plays an important role in the immune response (Müller and Rao, 2010). While the effect on NF-κB activity is less well investigated, preliminary data clearly indicate that Nef-mediated down-modulation of CD3 would also block activation of this transcription factor (**Figure 3C**). Thus, our current knowledge suggests that most primate lentiviruses may prevent NF-κB activation by Nef-mediated down-modulation of CD3, whereas HIV-1 and its simian precursors utilize Vpu to inhibit NF-κB signaling further downstream in the cascade. The former mechanism is associated with a more "resting" phenotype of virally infected T cells and disrupts their interaction with and responsiveness to other immune cells (Arhel et al., 2009). The latter may still allow activation of the infected T cells e.g., by APCs but prevent NF-κB-dependent antiviral gene expression during later stages of the viral replication cycle. Notably, most SIVs use their Nef protein to counteract tetherin (Jia et al., 2009; Sauter et al., 2009; Zhang et al., 2009). However, this does most likely not affect NF-κB activity since only the human but not monkey orthologs of tetherin are known to activate this transcription factor (Galão et al., 2012). Finally, HIV-2 uses Env to counteract restriction by human tetherin (Le Tortorec and Neil, 2009) but it has not been reported whether this mechanism also suppresses induction of NF-κB activity.

Importantly, the effects of all primate lentiviral Nefs on T cell activation may differ during the early and late stages of the viral replication cycle. Perhaps most notably, Nef proteins from HIV-2 and SIVs that down-modulate CD3 enhance IKKβ-induced NF-κB activation as efficiently as HIV-1 or SIV Nefs lacking the CD3 down-modulation function entirely (Sauter et al., 2015). Thus, primate lentiviral Nef proteins may generally boost the responsiveness to stimulation during the earliest stage of infection. While down-modulation of CD3 from the cell surface by Nef proteins possessing this function is highly effective, it may take more time than providing an initial boost to NF-κB activity to initiate viral gene expression and productive replication. Consequently, early stimulation and late inhibition of NF-κB may both be achieved by most primate lentiviruses and be mediated by cooperative Nef and Vpu functions in HIV-1 and its precursors. Finally, it is noteworthy that even HIV-1 Nef might have differential effects depending on the stage of infection. It is puzzling that for HIV-1 Nef enhancing (Herbein et al., 2008; Mangino et al., 2011), inhibitory (Niederman et al., 1992; Bandres and Ratner, 1994), and no (Yoon and Kim, 1999; Witte et al., 2008) effects on NF-κB activity have been reported. In part, these differences may depend on the state of activation of infected T cells. However, it will also be interesting to further examine whether the effect of HIV-1 Nef might differ at different stages of the viral replication cycle. In agreement with this possibility, it is known that HIV-1 Nefs down-modulate CD28, an important costimulatory factor of T cell activation, albeit substantially less effectively than HIV-2 and most SIV Nefs (Bell et al., 2001; Swigut et al., 2001). Timing adds another degree of complexity to the already complex role of the multi-functional Nef protein throughout the viral life cycle and clearly warrants further examination.

CONCLUSION AND PERSPECTIVES

Numerous studies have investigated how accessory proteins of HIV and SIV, i.e., Vif, Vpr, Nef, Vpu and/or Vpx, counteract antiviral restriction factors, such as APOBEC proteins, tetherin, SAMHD1 and SERINC3/5. Expression of most of these and other antiviral factors is induced by a very limited number of transcription factors that are activated upon viral immune sensing including NF-κB and IRFs. Modulation of

the induction and activity of these key regulators of innate and adaptive immunity may have a major benefit for viral replication but has only recently gained significant scientific attention. Specifically, recent data provide evidence that primate lentiviruses tightly regulate the activity of NF-κB to initiate efficient viral transcription while minimizing the expression of antiviral genes. As outlined above, the most common mechanism amongst primate lentiviruses may be initial boosting of NF-κB activity by Nef and prevention of further stimulation of T cell activation by potent down-modulation of TCR-CD3. This strategy seems to be highly successful considering the high prevalence of these viruses and the benign relationship with their natural simian hosts. In contrast, HIV-1 and its SIV precursors seem to use Nef to initially boost and Vpu to later on suppress NF-κB activation. A variety of Vpu functions that might provide a selective advance for viral immune evasion and replication have been reported (Sandberg et al., 2012; Sugden et al., 2016). Nonetheless, the emergence of a *vpu* containing subset of primate lentiviruses is somewhat surprising as they seem to be less prevalent and potentially more virulent in their natural hosts than other SIVs. Perhaps they have advantages in specific hosts as supported by the more efficient spread of HIV-1 in the human population compared to HIV-2 despite higher virulence. Notably, a few primate lentiviruses lack both *vpu* and the CD3 down-modulation function of Nef (Schmökel et al., 2011) and our preliminary data suggest that they might use yet another accessory protein (i.e., Vpr) to suppress NF-κB activity during the late stages of infection.

Altogether, recent evidence suggests that primate lentiviruses have evolved several sophisticated mechanisms to tightly regulate the activity of NF-κB and possibly other transcription factors throughout their replication cycle. The conservation of these functions supports an important role for viral replication and immune evasion. Furthermore, modulation of NF-κB activity is most likely also relevant for viral latency and inflammatory responses. NF-κB is targeted for the treatment of inflammatory and proliferative diseases as well as cancer and activation of the latent reservoirs of HIV-1 (Jiang and Dandekar, 2015; Panday et al., 2016). Because of its important role in many physiological processes, however, therapeutic modulation of NF-κB activity is prone to undesired adverse effects. Nevertheless, further studies on the mechanisms used by HIV-1 and other primate lentiviruses to manipulate this central transcription factor may provide important information on how these viruses establish latency and induce inflammation and perhaps even on how to prevent this.

AUTHOR CONTRIBUTIONS

EH and FK both wrote the manuscript and prepared figures.

FUNDING

FK is supported by the Deutsche Forschungsgemeinschaft (DFG) priority program "Innate Sensing and Restriction of Retroviruses" (SPP 1923) and an ERC Advanced grant "Anti-virome."

ACKNOWLEDGMENT

We thank Daniel Sauter, Dré van der Merwe and Dominik Hotter for comments and discussion.

REFERENCES

Akari, H., Bour, S., Kao, S., Adachi, A., and Strebel, K. (2001). The human immunodeficiency virus type 1 accessory protein Vpu induces apoptosis by suppressing the nuclear factor κB-dependent expression of antiapoptotic factors. *J. Exp. Med.* 194, 1299–1311. doi: 10.1084/jem.194.9.1299

Alexander, L., Du, Z., Rosenzweig, M., Jung, J. U., and Desrosiers, R. C. (1997). A role for natural simian immunodeficiency virus and human immunodeficiency virus type 1 nef alleles in lymphocyte activation. *J. Virol.* 71, 6094–6099.

Arhel, N., Lehmann, M., Clauss, K., Nienhaus, G. U., Piguet, V., and Kirchhoff, F. (2009). The inability to disrupt the immunological synapse between infected human T cells and APCs distinguishes HIV-1 from most other primate lentiviruses. *J. Clin. Invest.* 119, 2965–2975. doi: 10.1172/JCI 38994

Arhel, N. J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., et al. (2007). HIV-1 DNA flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.* 26, 3025–3037. doi: 10.1038/sj.emboj.7601740

Bachu, M., Yalla, S., Asokan, M., Verma, A., Neogi, U., Sharma, S., et al. (2012). Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. *J. Biol. Chem.* 287, 44714–44735. doi: 10.1074/jbc.M112. 397158

Bailes, E., Gao, F., Bibollet-Ruche, F., Courgnaud, V., Peeters, M., Marx, P. A., et al. (2003). Hybrid origin of SIV in Chimpanzees. *Science* 300, 1713–1713. doi: 10.1126/science.1080657

Bandres, J. C., and Ratner, L. (1994). Human immunodeficiency virus type 1 Nef protein down-regulates transcription factors NF-kappa B and AP-1 in human T cells in vitro after T-cell receptor stimulation. *J. Virol.* 68, 3243–3249.

Barboric, M., Nissen, R. M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B. M. (2001). NF-κB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* 8, 327–337. doi: 10.1016/S1097-2765(01)00314-8

Baur, A. S., Sass, G., Laffert, B., Willbold, D., Cheng-Mayer, C., and Peterlin, B. M. (1997). The N-terminus of Nef from HIV-1/SIV associates with a protein complex containing Lck and a serine kinase. *Immunity* 6, 283–291. doi: 10.1016/S1074-7613(00)80331-3

Baur, A. S., Sawai, E. T., Dazin, P., Fantl, W. J., Cheng-Mayer, C., and Matija Peterlin, B. (1994). HIV-1 nef leads to inhibition or activation of T cells depending on its intracellular localization. *Immunity* 1, 373–384. doi: 10.1016/1074-7613(94)90068-X

Bell, I., Ashman, C., Maughan, J., Hooker, E., Cook, F., and Reinhart, T. A. (1998). Association of simian immunodeficiency virus Nef with the T-cell receptor (TCR) ζ chain leads to TCR down-modulation. *J. Gen. Virol.* 79, 2717–2727. doi: 10.1099/0022-1317-79-11-2717

Bell, I., Schaefer, T. M., Trible, R. P., Amedee, A., and Reinhart, T. A. (2001). Down-modulation of the costimulatory molecule, CD28, is a conserved activity of multiple SIV Nefs and is dependent on histidine 196 of Nef. *Virology* 283, 148–158. doi: 10.1006/viro.2001.0872

Blanco-Melo, D., Venkatesh, S., and Bieniasz, P. D. (2016). Origins and evolution of tetherin, an orphan antiviral gene. *Cell Host Microbe* 20, 189–201. doi: 10.1016/j.chom.2016.06.007

Boomer, J. S., and Green, J. M. (2010). An enigmatic tail of CD28 signaling. *Cold Spring Harb. Perspect. Biol.* 2:a002436. doi: 10.1101/cshperspect.a002436

Bour, S., Perrin, C., Akari, H., and Strebel, K. (2001). The human immunodeficiency virus type 1 Vpu protein inhibits NF-κB activation by interfering with βTrCP-mediated degradation of IκB. *J. Biol. Chem.* 276, 15920–15928. doi: 10.1074/jbc.M010533200

Cary, D. C., Fujinaga, K., and Peterlin, B. M. (2016). Molecular mechanisms of HIV latency. *J. Clin. Invest.* 126, 448–454. doi: 10.1172/JCI80565

Chahroudi, A., Bosinger, S. E., Vanderford, T. H., Paiardini, M., and Silvestri, G. (2012). Natural SIV hosts: showing AIDS the door. *Science* 335, 1188–1193. doi: 10.1126/science.1217550

Chan, J. K., and Greene, W. C. (2012). Dynamic roles for NF-κB in HTLV-I and HIV-1 retroviral pathogenesis. *Immunol. Rev.* 246, 286–310. doi: 10.1111/j.1600-065X.2012.01094.x

Cheng, J., Montecalvo, A., and Kane, L. P. (2011). Regulation of NF-κB induction by TCR/CD28. *Immunol. Res.* 50, 113–117. doi: 10.1007/s12026-011-8216-z

Cildir, G., Low, K. C., and Tergaonkar, V. (2016). Noncanonical NF-κB signaling in health and disease. *Trends Mol. Med.* 22, 414–429. doi: 10.1016/j.molmed.2016.03.002

Cocka, L. J., and Bates, P. (2012). Identification of alternatively translated Tetherin isoforms with differing antiviral and signaling activities. *PLoS Pathog.* 8:e1002931. doi: 10.1371/journal.ppat.1002931

Colomer-Lluch, M., Gollahon, L. S., and Serra-Moreno, R. (2016). Anti-HIV factors: targeting each step of HIV's replication cycle. *Curr. HIV Res.* 14, 175–182. doi: 10.2174/1570162X14999160224094621

Coudronniere, N., Villalba, M., Englund, N., and Altman, A. (2000). NF-κB activation induced by T cell receptor/CD28 costimulation is mediated by protein kinase C-θ. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3394–3399. doi: 10.1073/pnas.97.7.3394

Dharan, A., Talley, S., Tripathi, A., Mamede, J. I., Majetschak, M., Hope, T. J., et al. (2016). KIF5B and Nup358 cooperatively mediate the nuclear import of HIV-1 during infection. *PLOS Pathog.* 12:e1005700. doi: 10.1371/journal.ppat.1005700

Doyle, T., Goujon, C., and Malim, M. H. (2015). HIV-1 and interferons: who's interfering with whom? *Nat. Rev. Microbiol.* 13, 403–413. doi: 10.1038/nrmicro3449

Fenard, D., Yonemoto, W., de Noronha, C., Cavrois, M., Williams, S. A., and Greene, W. C. (2005). Nef is physically recruited into the immunological synapse and potentiates T cell activation early after TCR engagement. *J. Immunol.* 175, 6050–6057. doi: 10.4049/JIMMUNOL.175.9.6050

Fiume, G., Vecchio, E., De Laurentiis, A., Trimboli, F., Palmieri, C., Pisano, A., et al. (2012). Human immunodeficiency virus-1 Tat activates NF-κB via physical interaction with IκB-α and p65. *Nucleic Acids Res.* 40, 3548–3562. doi: 10.1093/nar/gkr1224

Fortin, J.-F., Barat, C., Beauséjour, Y., Barbeau, B., and Tremblay, M. J. (2004). Hyper-responsiveness to stimulation of human immunodeficiency virus-infected $CD4^+$ T cells requires Nef and Tat virus gene products and results from higher NFAT, NF-κB, and AP-1 induction. *J. Biol. Chem.* 279, 39520–39531. doi: 10.1074/jbc.M407477200

Galão, R. P., Le Tortorec, A., Pickering, S., Kueck, T., and Neil, S. J. D. (2012). Innate sensing of HIV-1 assembly by Tetherin induces NFκB-dependent proinflammatory responses. *Cell Host Microbe* 12, 633–644. doi: 10.1016/j.chom.2012.10.007

Galão, R. P., Pickering, S., Curnock, R., and Neil, S. J. D. (2014). Retroviral retention activates a syk-dependent HemITAM in human tetherin. *Cell Host Microbe* 16, 291–303. doi: 10.1016/j.chom.2014.08.005

Ghosh, S., and Hayden, M. S. (2012). Celebrating 25 years of NF-κB research. *Immunol. Rev.* 246, 5–13. doi: 10.1111/j.1600-065X.2012.01111.x

Hayden, M. S., West, A. P., and Ghosh, S. (2006). SnapShot: NF-κB signaling pathways. *Cell* 127, 1286–1287. doi: 10.1016/j.cell.2006.12.005

Heigele, A., Kmiec, D., Regensburger, K., Langer, S., Peiffer, L., Stürzel, C. M., et al. (2016). The potency of nef-mediated SERINC5 antagonism correlates with the prevalence of primate lentiviruses in the wild. *Cell Host Microbe* 20, 381–391. doi: 10.1016/j.chom.2016.08.004

Herbein, G., Varin, A., Larbi, A., Fortin, C., Mahlknecht, U., Fulop, T., et al. (2008). Nef and TNFα are coplayers that favor HIV-1 replication in monocytic cells and primary macrophages. *Curr. HIV Res.* 6, 117–129. doi: 10.2174/157016208783884985

Heusinger, E., Kluge, S. F., Kirchhoff, F., and Sauter, D. (2015). Early vertebrate evolution of the host restriction factor tetherin. *J. Virol.* 89, 12154–12165. doi: 10.1128/JVI.02149-15

Hiscott, J., Kwon, H., Génin, P., Ghosh, S., May, M., Kopp, E., et al. (2001). Hostile takeovers: viral appropriation of the NF-κB pathway. *J. Clin. Invest.* 107, 143–151. doi: 10.1172/JCI11918

Hotter, D., Sauter, D., and Kirchhoff, F. (2013). Emerging role of the host restriction factor tetherin in viral immune sensing. *J. Mol. Biol.* 425, 4956–4964. doi: 10.1016/j.jmb.2013.09.029

Iafrate, A. J., Bronson, S., and Skowronski, J. (1997). Separable functions of Nef disrupt two aspects of T cell receptor machinery: CD4 expression and CD3 signaling. *EMBO J.* 16, 673–684. doi: 10.1093/emboj/16.4.673

Jakobsen, M. R., Olagnier, D., and Hiscott, J. (2015). Innate immune sensing of HIV-1 infection. *Curr. Opin. HIV AIDS* 10, 96–102. doi: 10.1097/COH.0000000000000129

Janardhan, A., Swigut, T., Hill, B., Myers, M. P., and Skowronski, J. (2004). HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis. *PLoS Biol.* 2:E6. doi: 10.1371/journal.pbio.0020006

Jia, B., Serra-Moreno, R., Neidermyer, W., Rahmberg, A., Mackey, J., Fofana, I. B., et al. (2009). Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by Tetherin/BST2. *PLoS Pathog.* 5:e1000429. doi: 10.1371/journal.ppat.1000429

Jiang, G., and Dandekar, S. (2015). Targeting NF-κB signaling with protein kinase C agonists as an emerging strategy for combating HIV latency. *AIDS Res. Hum. Retroviruses* 31, 4–12. doi: 10.1089/AID.2014.0199

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-κB activity. *Annu. Rev. Immunol.* 18, 621–663. doi: 10.1146/annurev.immunol.18.1.621

Keele, B. F., Jones, J. H., Terio, K. A., Estes, J. D., Rudicell, R. S., Wilson, M. L., et al. (2009). Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* 460, 515–519. doi: 10.1038/nature08200

Khalid, M., Yu, H., Sauter, D., Usmani, S. M., Schmokel, J., Feldman, J., et al. (2012). Efficient Nef-mediated downmodulation of TCR-CD3 and CD28 is associated with high $CD4^+$ T cell counts in viremic HIV-2 infection. *J. Virol.* 86, 4906–4920. doi: 10.1128/JVI.06856-11

Kinoshita, S., Chen, B. K., Kaneshima, H., and Nolan, G. P. (1998). Host control of HIV-1 parasitism in T Cells by the nuclear factor of activated T cells. *Cell* 95, 595–604. doi: 10.1016/S0092-8674(00)81630-X

Kirchhoff, F. (2009). Is the high virulence of HIV-1 an unfortunate coincidence of primate lentiviral evolution? *Nat. Rev. Microbiol.* 7:467. doi: 10.1038/nrmicro2111

Kirchhoff, F. (2010). Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. *Cell Host Microbe* 8, 55–67. doi: 10.1016/j.chom.2010.06.004

Kluge, S. F., Sauter, D., and Kirchhoff, F. (2015). SnapShot: antiviral restriction factors. *Cell* 163, 774–774.e1. doi: 10.1016/j.cell.2015.10.019

Kogan, M., Deshmane, S., Sawaya, B. E., Gracely, E. J., Khalili, K., and Rappaport, J. (2013). Inhibition of NF-κB activity by HIV-1 Vpr is dependent on Vpr binding protein. *J. Cell. Physiol.* 228, 781–790. doi: 10.1002/jcp.24226

Kumar, A., Abbas, W., Colin, L., Khan, K. A., Bouchat, S., Varin, A., et al. (2016). Tuning of AKT-pathway by Nef and its blockade by protease inhibitors results in limited recovery in latently HIV infected T-cell line. *Sci. Rep.* 6:24090. doi: 10.1038/srep24090

Le Tortorec, A., and Neil, S. J. D. (2009). Antagonism to and intracellular sequestration of human tetherin by the human immunodeficiency virus type 2 envelope glycoprotein. *J. Virol.* 83, 11966–11978. doi: 10.1128/JVI.01515-09

Lelek, M., Di Nunzio, F., Henriques, R., Charneau, P., Arhel, N., and Zimmer, C. (2012). Superresolution imaging of HIV in infected cells with FlAsH-PALM. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8564–8569. doi: 10.1073/pnas.1013267109

Liang, Z., Liu, R., Lin, Y., Liang, C., Tan, J., and Qiao, W. (2015). HIV-1 Vpr protein activates the NF-κB pathway to promote G2/M cell cycle arrest. *Virol. Sin.* 30, 441–448. doi: 10.1007/s12250-015-3654-8

Liu, R., Lin, Y., Jia, R., Geng, Y., Liang, C., Tan, J., et al. (2014). HIV-1 Vpr stimulates NF-κB and AP-1 signaling by activating TAK1. *Retrovirology* 11:45. doi: 10.1186/1742-4690-11-45

Liu, R., Tan, J., Lin, Y., Jia, R., Yang, W., Liang, C., et al. (2013). HIV-1 Vpr activates both canonical and noncanonical NF-κB pathway by enhancing the phosphorylation of IKKα/β. *Virology* 439, 47–56. doi: 10.1016/j.virol.2013.01.020

Mangino, G., Percario, Z. A., Fiorucci, G., Vaccari, G., Acconcia, F., Chiarabelli, C., et al. (2011). HIV-1 Nef induces proinflammatory state in macrophages through its acidic cluster domain: involvement of TNF alpha receptor associated factor 2. *PLoS ONE* 6:e22982. doi: 10.1371/journal.pone.0022982

McCool, K. W., and Miyamoto, S. (2012). DNA damage-dependent NF-κB activation: NEMO turns nuclear signaling inside out. *Immunol. Rev.* 246, 311–326. doi: 10.1111/j.1600-065X.2012.01101.x

Miyamoto, S. (2011). Nuclear initiated NF-κB signaling: NEMO and ATM take center stage. *Cell Res.* 21, 116–130. doi: 10.1038/cr.2010.179

Müller, M. R., and Rao, A. (2010). NFAT, immunity and cancer: a transcription factor comes of age. *Nat. Rev. Immunol.* 10, 645–656. doi: 10.1038/nri2818

Nabel, G., and Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 329, 219–222.

Nabel, G. J., and Verma, I. M. (1993). Proposed NF-κB/IκB family nomenclature. *Genes Dev.* 7, 2063. doi: 10.1101/gad.7.11.2063

Napetschnig, J., and Wu, H. (2013). Molecular basis of NF-κB signaling. *Annu. Rev. Biophys.* 42, 443–468. doi: 10.1146/annurev-biophys-083012-130338

Nelson, D. E., Ihekwaba, A. E. C., Elliott, M., Johnson, J. R., Kell, D. B., and White, M. R. H. (2004). Oscillations in NF-κB signaling control the dynamics of gene expression. *Science* 306, 704–709. doi: 10.1126/science.1099962

Niederman, T. M., Garcia, J. V., Hastings, W. R., Luria, S., and Ratner, L. (1992). Human immunodeficiency virus type 1 Nef protein inhibits NF-κB induction in human T cells. *J. Virol.* 66, 6213–6219.

Nyamweya, S., Hegedus, A., Jaye, A., Rowland-Jones, S., Flanagan, K. L., and Macallan, D. C. (2013). Comparing HIV-1 and HIV-2 infection: lessons for viral immunopathogenesis. *Rev. Med. Virol.* 23, 221–240. doi: 10.1002/rmv.1739

Oeckinghaus, A., and Ghosh, S. (2009). The NF-κB family of transcription factors and its regulation. *Cold Spring Harb. Perspect. Biol.* 1:a000034. doi: 10.1101/cshperspect.a000034

Panday, A., Inda, M. E., Bagam, P., Sahoo, M. K., Osorio, D., and Batra, S. (2016). Transcription factor NF-κB: an update on intervention strategies. *Arch. Immunol. Ther. Exp. (Warsz)* 64, 463–483. doi: 10.1007/s00005-016-0405-y

Paul, S., and Schaefer, B. C. (2013). A new look at T cell receptor signaling to nuclear factor-κB. *Trends Immunol.* 34, 269–281. doi: 10.1016/j.it.2013.02.002

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., et al. (2011). TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* 472, 361–365. doi: 10.1038/nature09976

Pfeffer, L. M. (2011). The role of nuclear factor κB in the interferon response. *J. Interf. cytokine Res.* 31, 553–559. doi: 10.1089/jir.2011.0028

Postler, T. S., and Desrosiers, R. C. (2012). The cytoplasmic domain of the HIV-1 glycoprotein gp41 induces NF-κB activation through TGF-β-activated kinase 1. *Cell Host Microbe* 11, 181–193. doi: 10.1016/j.chom.2011.12.005

Rasaiyaah, J., Tan, C. P., Fletcher, A. J., Price, A. J., Blondeau, C., Hilditch, L., et al. (2013). HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* 503, 402–405. doi: 10.1038/nature12769

Roux, P., Alfieri, C., Hrimech, M., Cohen, E. A., and Tanner, J. E. (2000). Activation of transcription factors NF-κB and NF-IL-6 by human immunodeficiency virus type 1 protein R (Vpr) induces interleukin-8 expression. *J. Virol.* 74, 4658–4665. doi: 10.1128/JVI.74.10.4658-4665.2000

Sandberg, J. K., Andersson, S. K., Bächle, S. M., Nixon, D. F., and Moll, M. (2012). HIV-1 Vpu interference with innate cell-mediated immune mechanisms. *Curr. HIV Res.* 10, 327–333. doi: 10.2174/157016212800792513

Sauter, D., Hotter, D., Van Driessche, B., Stürzel, C. M., Kluge, S. F., Wildum, S., et al. (2015). Differential regulation of NF-κB-mediated proviral and antiviral host gene expression by primate lentiviral nef and vpu proteins. *Cell Rep.* 10, 586–600. doi: 10.1016/j.celrep.2014.12.047

Sauter, D., and Kirchhoff, F. (2016). IFITMs: important factors in transmission of HIV-1. *Cell Host Microbe* 20, 407–408. doi: 10.1016/j.chom.2016.09.009

Sauter, D., Schindler, M., Specht, A., Landford, W. N., Münch, J., Kim, K.-A., et al. (2009). Tetherin-Driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. *Cell Host Microbe* 6, 409–421. doi: 10.1016/j.chom.2009.10.004

Sawai, E. T., Baur, A., Struble, H., Peterlin, B. M., Levy, J. A., and Cheng-Mayer, C. (1994). Human immunodeficiency virus type 1 Nef associates with a cellular serine kinase in T lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1539–1543. doi: 10.1073/pnas.91.4.1539

Schindler, M., Münch, J., Kutsch, O., Li, H., Santiago, M. L., Bibollet-Ruche, F., et al. (2006). Nef-mediated suppression of T cell activation was lost in a lentiviral lineage that gave rise to HIV-1. *Cell* 125, 1055–1067. doi: 10.1016/j.cell.2006.04.033

Schindler, M., Schmökel, J., Specht, A., Li, H., Münch, J., Khalid, M., et al. (2008). Inefficient Nef-mediated downmodulation of CD3 and MHC-I correlates with loss of $CD4^+$T cells in natural SIV infection. *PLoS Pathog.* 4:e1000107. doi: 10.1371/journal.ppat.1000107

Schmökel, J., Sauter, D., Schindler, M., Leendertz, F. H., Bailes, E., Dazza, M.-C., et al. (2011). The presence of a vpu gene and the lack of Nef-mediated downmodulation of T cell receptor-CD3 are not always linked in primate lentiviruses. *J. Virol.* 85, 742–752. doi: 10.1128/JVI.02087-10

Schrager, J. A., Der Minassian, V., and Marsh, J. W. (2002). HIV Nef increases T cell ERK MAP kinase activity. *J. Biol. Chem.* 277, 6137–6142. doi: 10.1074/jbc.M107322200

Schrager, J. A., and Marsh, J. W. (1999). HIV-1 Nef increases T cell activation in a stimulus-dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* 96, 8167–8172. doi: 10.1073/PNAS.96.14.8167

Sen, R., and Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* 46, 705–716. doi: 10.1016/0092-8674(86)90346-6

Simmons, A., Aluvihare, V., and McMichael, A. (2001). Nef triggers a transcriptional program in T cells imitating single-signal T cell activation and inducing HIV virulence mediators. *Immunity* 14, 763–777. doi: 10.1016/S1074-7613(01)00158-3

Skaug, B., Jiang, X., and Chen, Z. J. (2009). The role of ubiquitin in NF-κB regulatory pathways. *Annu. Rev. Biochem.* 78, 769–796. doi: 10.1146/annurev.biochem.78.070907.102750

Sloan, R. D., Wainberg, M. A., Bar-Magen, T., Sloan, R., Faltenbacher, V., Donahue, D., et al. (2011). The role of unintegrated DNA in HIV infection. *Retrovirology* 8:52. doi: 10.1186/1742-4690-8-52

Sodora, D. L., Allan, J. S., Apetrei, C., Brenchley, J. M., Douek, D. C., Else, J. G., et al. (2009). Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. *Nat. Med.* 15, 861–865. doi: 10.1038/nm.2013

Sugden, S. M., Bego, M. G., Pham, T. N. Q., and Cohen, É. A. (2016). Remodeling of the host cell plasma membrane by HIV-1 Nef and Vpu: a strategy to ensure viral fitness and persistence. *Viruses* 8:67. doi: 10.3390/v8030067

Swigut, T., Shohdy, N., and Skowronski, J. (2001). Mechanism for down-regulation of CD28 by Nef. *EMBO J.* 20, 1593–1604. doi: 10.1093/emboj/20.7.1593

Takeuchi, J. S., Ren, F., Yoshikawa, R., Yamada, E., Nakano, Y., Kobayashi, T., et al. (2015). Coevolutionary dynamics between tribe Cercopithecini tetherins and their lentiviruses. *Sci. Rep.* 5:16021. doi: 10.1038/srep16021

Tokarev, A., Suarez, M., Kwan, W., Fitzpatrick, K., Singh, R., and Guatelli, J. (2013). Stimulation of NF-κB activity by the HIV restriction factor BST2. *J. Virol.* 87, 2046–2057. doi: 10.1128/JVI.02272-12

Tong-Starksen, S. E., Luciw, P. A., and Peterlin, B. M. (1987). Human immunodeficiency virus long terminal repeat responds to T-cell activation signals. *Proc. Natl. Acad. Sci. U.S.A.* 84, 6845–6849. doi: 10.1073/pnas.84.19.6845

Utay, N. S., and Douek, D. C. (2016). Interferons and HIV infection: the good, the bad, and the Ugly. *Pathog. Immun.* 1, 107–116. doi: 10.20411/pai.v1i1.125

Vallabhapurapu, S., and Karin, M. (2009). Regulation and function of NF-κB transcription factors in the immune system. *Annu. Rev. Immunol.* 27, 693–733. doi: 10.1146/annurev.immunol.021908.132641

Varin, A., Decrion, A.-Z., Sabbah, E., Quivy, V., Sire, J., Van Lint, C., et al. (2005). Synthetic Vpr protein activates activator protein-1, c-Jun N-terminal kinase, and NF-κB and stimulates HIV-1 transcription in promonocytic cells and primary macrophages. *J. Biol. Chem.* 280, 42557–42567. doi: 10.1074/jbc.M502211200

Wang, D., and Baldwin, A. S. J. (1998). Activation of nuclear factor-κB-dependent transcription by tumor necrosis factor-α is mediated through phosphorylation of RelA/p65 on serine 529. *J. Biol. Chem.* 273, 29411–29416. doi: 10.1074/jbc.273.45.29411

Williams, S. A., Chen, L.-F., Kwon, H., Fenard, D., Bisgrove, D., Verdin, E., et al. (2004). Prostratin antagonizes HIV latency by activating NF-κB. *J. Biol. Chem.* 279, 42008–42017. doi: 10.1074/jbc.M402124200

Williams, S. A., Chen, L.-F., Kwon, H., Ruiz-Jarabo, C. M., Verdin, E., and Greene, W. C. (2006). NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO J.* 25, 139–149. doi: 10.1038/sj.emboj.7600900

Williams, S. A., Kwon, H., Chen, L.-F., and Greene, W. C. (2007). Sustained induction of NF-κB is required for efficient expression of latent human immunodeficiency virus type 1. *J. Virol.* 81, 6043–6056. doi: 10.1128/JVI.02074-06

Witte, V., Laffert, B., Gintschel, P., Krautkrämer, E., Blume, K., Fackler, O. T., et al. (2008). Induction of HIV transcription by Nef involves Lck activation and protein kinase C theta raft recruitment leading to activation of ERK1/2 but not NF kappa B. *J. Immunol.* 181, 8425–8432. doi: 10.4049/jimmunol.181.12.8425

Xu, X. N., Laffert, B., Screaton, G. R., Kraft, M., Wolf, D., Kolanus, W., et al. (1999). Induction of Fas ligand expression by HIV involves the interaction of Nef with the T cell receptor ζ chain. *J. Exp. Med.* 189, 1489–1496. doi: 10.1084/jem.189.9.1489

Yoon, K., and Kim, S. (1999). Lack of negative influence on the cellular transcription factors NF-κB and AP-1 by the nef protein of human immunodeficiency virus type 1. *J. Gen. Virol.* 80, 2951–2956. doi: 10.1099/0022-1317-80-11-2951

Yu, H., Khalid, M., Heigele, A., Schmökel, J., Usmani, S. M., van der Merwe, J., et al. (2015). Lentiviral Nef proteins manipulate T cells in a subset-specific manner. *J. Virol.* 89, 1986–2001. doi: 10.1128/JVI.03104-14

Zhang, F., Wilson, S. J., Landford, W. C., Virgen, B., Gregory, D., Johnson, M. C., et al. (2009). Nef Proteins from simian immunodeficiency viruses are tetherin antagonists. *Cell Host Microbe* 6, 54–67. doi: 10.1016/j.chom.2009.05.008

Zhang, L., Huang, Y., Yuan, H., Chen, B. K., Ip, J., and Ho, D. D. (1997). Identification of a replication-competent pathogenic human immunodeficiency virus type 1 with a duplication in the TCF-1α region but lacking NF-κB binding sites. *J. Virol.* 71, 1651–1656.

Zhong, H., May, M. J., Jimi, E., and Ghosh, S. (2002). The phosphorylation status of nuclear NF-κB determines its association with CBP/p300 or HDAC-1. *Mol. Cell* 9, 625–636. doi: 10.1016/S1097-2765(02)00477-X

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Heusinger and Kirchhoff. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

MINI REVIEW
published: 30 January 2017
doi: 10.3389/fmicb.2017.00080

HIV-1 Tat and Viral Latency: What We Can Learn from Naturally Occurring Sequence Variations

*Doreen Kamori[1] and Takamasa Ueno[1,2]**

[1] Center for AIDS Research, Kumamoto University, Kumamoto, Japan, [2] International Research Center for Medical Sciences, Kumamoto University, Kumamoto, Japan

Despite the effective use of antiretroviral therapy, the remainder of a latently HIV-1-infected reservoir mainly in the resting memory $CD4^+$ T lymphocyte subset has provided a great setback toward viral eradication. While host transcriptional silencing machinery is thought to play a dominant role in HIV-1 latency, HIV-1 protein such as Tat, may affect both the establishment and the reversal of latency. Indeed, mutational studies have demonstrated that insufficient Tat transactivation activity can result in impaired transcription of viral genes and the establishment of latency in cell culture experiments. Because Tat protein is one of highly variable proteins within HIV-1 proteome, it is conceivable that naturally occurring Tat mutations may differentially modulate Tat functions, thereby influencing the establishment and/or the reversal of viral latency *in vivo*. In this mini review, we summarize the recent findings of Tat naturally occurring polymorphisms associating with host immune responses and we highlight the implication of Tat sequence variations in relation to HIV latency.

Keywords: HIV-1, Tat, latency, transactivation, variability, reactivation

OPEN ACCESS

Edited by:
Hirofumi Akari,
Kyoto University, Japan

Reviewed by:
Kazuhisa Yoshimura,
National Institute of Infectious Diseases, Japan
Taketoshi Mizutani,
Institute of Microbial Chemistry, Japan

****Correspondence:***
Takamasa Ueno
uenotaka@kumamoto-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *06 December 2016*
Accepted: *11 January 2017*
Published: *30 January 2017*

Citation:
Kamori D and Ueno T (2017) HIV-1 Tat and Viral Latency: What We Can Learn from Naturally Occurring Sequence Variations. Front. Microbiol. 8:80. doi: 10.3389/fmicb.2017.00080

INTRODUCTION

Viral latency is a reversible state whereby a pathogenic virus becomes dormant (latent) during the viral life cycle in individual cells. HIV-1 may either actively replicate to rapidly produce progeny virions or can enter a long-lived quiescent state (viral latency), from which it may later be subsequently reactivated. The mechanisms for establishment and maintenance of HIV-1 latency mainly operate at the transcriptional level by both viral (Yukl et al., 2009; Donahue et al., 2012; Donahue and Wainberg, 2013; Ranasinghe et al., 2013) and host (Coiras et al., 2009, 2010; Donahue and Wainberg, 2013) machineries and occur at the levels of transcription, chromatin modification, and epigenetic regulations (Coiras et al., 2009; Donahue and Wainberg, 2013; Archin et al., 2014; Cary et al., 2016).

HIV-1 latency is primarily found within resting memory $CD4^+$ T cells (Chun et al., 1995, 1997; Dahabieh et al., 2015), microglia cells (Chakrabarti et al., 1991; Davis et al., 1992), monocytes/macrophages (Battistini and Sgarbanti, 2014; Kumar et al., 2014; Abbas et al., 2015), and others (Canki et al., 2001; MacDougall et al., 2002; Valentin et al., 2002) which intrinsically have a long half-life *in vivo*. Because the expression level of the viral proteins is absent or poorly expressed and also the existence of immune escape mutations (Deng et al., 2015), the latently infected cells are much less susceptible to be recognized and cleared by the host immune system, viral cytopathic effects or currently available antiretroviral drugs. Thus to date, latently infected viral reservoir is one of the fundamental limitations toward HIV cure (Marsden and Zack, 2015).

Among the viral proteins, HIV-1 Tat has attracted more attention in viral latency because it potently plays a role in viral transcription regulation. Structurally, Tat is a small nuclear protein with amino acid length ranging from 86 to 101 and the molecular weight ranging from 14 to 16 kDa (Ruben et al., 1989). Functionally, Tat is divided by six domains and plays a role in nuclear translocation (Efthymiadis et al., 1998; Rana and Jeang, 1999), binding for viral RNA (Roy et al., 1990), several host factors and co-factors (Jeang et al., 1993; Garber et al., 1998; Marzio et al., 1998), and the transactivation of 5′ long terminal repeat (LTR) (Ruben et al., 1989; Roy et al., 1990; Jeang et al., 1993; Tong-Starksen et al., 1993; Neuveut and Jeang, 1996). Despite such fundamental functions in the virus life cycle, Tat is a highly polymorphic protein comparable to other HIV-1 polymorphic proteins such as Env, Vpu, and Nef (Yusim et al., 2002; Rossenkhan et al., 2012). Recent studies indicate that a substantial part of viral polymorphisms including in Tat is caused by viral mutational escape from cellular immune responses (Allen et al., 2000; Mason et al., 2009; John et al., 2010; Carlson et al., 2012). It is conceivable that naturally occurring mutations in Tat may modulate transactivation or other Tat functions, and that consequently affect the establishment and reversal of HIV-1 latency *in vivo*. In this mini review, we will describe the role of HIV-1 Tat toward HIV-1 latency establishment and reactivation, and discuss the possibility that naturally occurring Tat mutations may influence viral latency. The details of host machinery in relation to HIV-1 latency have been well described in recent reviews (Ruelas and Greene, 2013; Dahabieh et al., 2015; Cary et al., 2016) and are not discussed here.

The Role of HIV-1 Tat in Establishment of Viral Latency

Tat ensures high levels of viral transcription during the virus life cycle (Das et al., 2011). The protein stimulates transcription from the viral 5′ LTR promoter and controls RNA polymerase II (RNAP II) elongation. This is achieved by Tat binding to the TAR hairpin in the nascent RNA transcript and the complex of positive transcription elongation factor b (P-TEFb) composed of Cyclin T1 (CycT1) and cyclin-dependent kinase 9 (CDK9) which phosphorylates the C-terminal domain of the RNAP II that consequently promote transcriptional elongation from the viral promoter (**Figure 1**) (Dahmus, 1996; Parada and Roeder, 1996; Das et al., 2011; Peterlin et al., 2012). Importantly, the absence or inactivation of Tat in HIV-1 infection has been observed to predominantly generate short non-polyadenylated transcripts of less than 100 nucleotides in length that forms the TAR stem-loop structure, and resulted in reduction of viral transcription and replication (Feng and Holland, 1988; Roy et al., 1990; Yedavalli et al., 2003; Pagans et al., 2005; Das et al., 2011) (**Figure 1**).

It could be therapeutically beneficial if we could prevent or at least reduce to a large extent the size of the established latent reservoir. Evidence indicates that Tat, when present in sufficient quantities, may counteract the establishment of HIV-1 latency by promoting transcriptional initiation or elongation (Pearson et al., 2008; Donahue et al., 2012). One study demonstrated that fewer latently infected cells were established in Jurkat cells that stably express Tat compared to cells that did not express Tat (Donahue et al., 2012). These findings highlight the contribution of Tat and its abundance on prevention of establishment of viral latency. In contrast, a complete block of Tat activity may induce permanent latency as observed with use the of Tat dependent transcription inhibitors such as didehydro-cortistatin A (dCA). The agent has been shown to induce permanently the inactivation of the viral transcription in primary latently infected $CD4^+$ T cells isolated from aviremic ART-treated subjects; and also when tested in several cell line models of latency (HeLa-CD4, promyelocytic OM-10.1 and J-Lat T-lymphocytic cell lines) (Mousseau et al., 2015). In addition, in the same study both in primary cells and latently infected cell line models, the dCA established a state of latency with an extremely impaired ability to reactivate even in the presence of conventional latency-reversing agents (such as TNF-α and prostratin). Therefore, the concomitant treatment of dCA and antiretroviral drugs may reduce the size of reactivation of latently infected cells *in vivo* and eventually attain a functional HIV cure. However, to date, most experiments done for dCA are limited to *in vitro* models of latently infected cell lines and primary $CD4^+$ T cells. Therefore, further studies are needed to test the efficacy and safety of dCA as a viral transcription inhibitor agent in advanced experimental systems such as using humanized mice and non-human primates.

Role of Tat Protein on Reversion of Viral Latency

Tat can also contribute to reactivation of latently infected cells. For example, previous studies demonstrated that Tat is responsible for directly activating viral transcription in the patient-derived latently infected resting memory $CD4^+$ T cells without requiring cellular activation (Lin et al., 2003; Lassen et al., 2006). This is also supported by the Jurkat model of latency showing that the introduction of exogenous Tat was sufficient to reactivate most of the latently infected population (Donahue et al., 2012). Similarly, HIV-1 latently infected cells, at least in Jurkat cells, can be reactivated by cellular superinfection in a Tat-dependent manner (Donahue et al., 2013). Moreover, both experimental and computational methods have revealed that Tat is more effective than cellular activation approaches in reactivation of full-length transcription of latent HIV. In a recent study, Razooky et al. (2015) showed that removal of cell activation stimuli in HIV-infected primary $CD4^+$ T cells resulted in a drastic decline in cellular activation, but viral transcription activity as measured by GFP expression of productively infected cells remained relatively unchanged. Furthermore, the same study revealed by a computational method of HIV transcriptional modulation that Tat in abundance alone is sufficient for reactivation of the latently infected cells (Razooky et al., 2015). In addition, the depletion of some host factors or molecules that inhibit Tat transactivation activities, such as the long non-coding RNAs (NRON) that degrades Tat protein, in combination with a histone deacetylase (HDAC) inhibitor, has also been shown to significantly reactivate HIV-1 latency in $CD4^+$ T lymphocytes (Li et al., 2016). Furthermore,

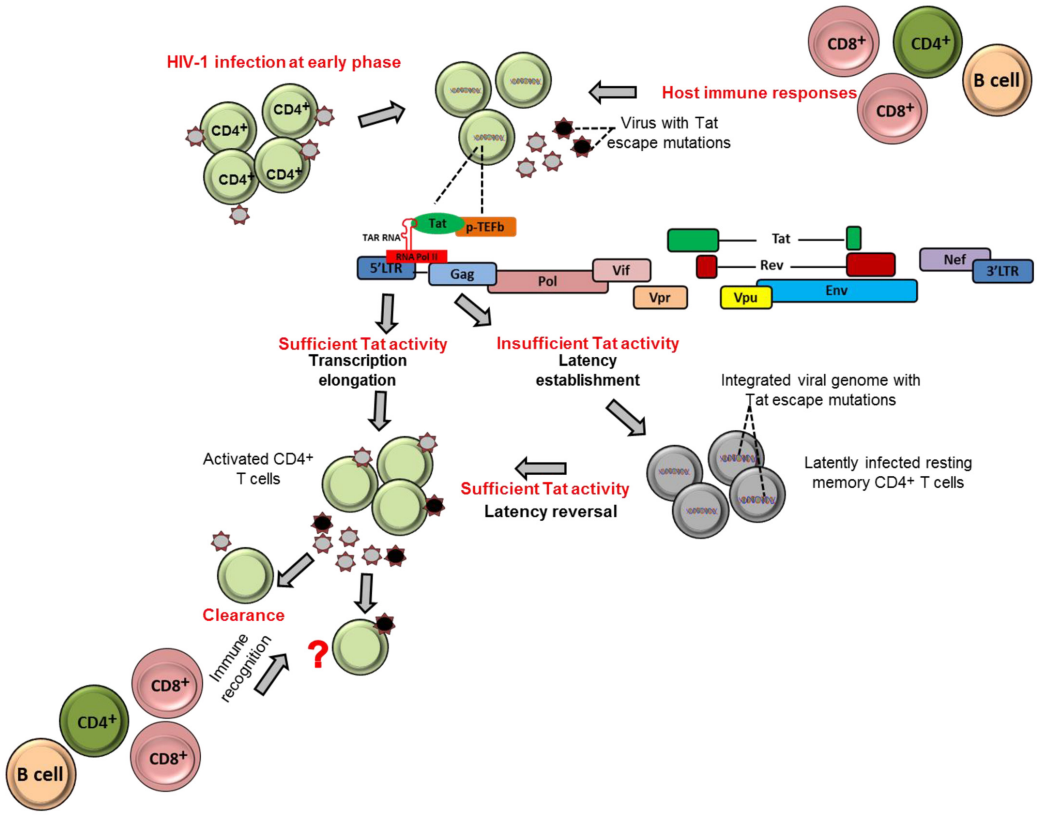


FIGURE 1 | Tat role in establishment and reversion of viral latency. The schematic figures illustrate how HIV-1 Tat may contribute to establishment of viral latency and latency reversal in resting memory CD4$^+$ T cells in HIV-1 infection.

in a recent mutational study, a Tat mutant, Tat-R5M4 that comprises of V36A, Q66A, V67A, S66A, and S77A mutations, exhibited a potent ability to reactivate latently infected CD4$^+$ T lymphocytes (Geng et al., 2016). Taken together, these findings provide a potential alternative approach toward reactivation of the latently infected cells with Tat protein.

Effects of Tat Variability on Latency

Sequence analysis of plasma viral RNA isolated from cross-sectional and longitudinal collection of HIV-infected individuals exhibited that HIV-1 Tat is a highly variable protein even among the rapidly mutating HIV-1 proteins such as Env, Vpu, and Nef (Yusim et al., 2002; Li et al., 2015). The high genetic variability of HIV-1 Tat is observed across the subtypes, such as subtypes B and C, in the major HIV-1 group M, and also across HIV-1 groups O and N as well as HIV-2 (Yusim et al., 2002; Rossenkhan et al., 2012; Li et al., 2015; Roy et al., 2015b). Interestingly, Bayesian evolutionary analysis model demonstrated that subtype B Tat has evolved relatively faster than other subtypes (Roy et al., 2015a). The extent of amino acid variability in Tat as estimated by the Shannon entropy score in subtype B sequences published in Los Alamos sequence database is illustrated in **Figure 2**.

Mutational studies of HIV-1 Tat revealed that Tat is divided into six functional domains (Kuppuswamy et al., 1989) (**Figure 2**). The first three domains are responsible for Tat transactivation activity and binding with the transcription cofactors (Feng and Holland, 1988; Feinberg et al., 1991; Garber et al., 1998; Wei et al., 1998; Rusnati et al., 1999); while the fourth domain is a TAR binding domain (Dingwall et al., 1989; Roy et al., 1990; Weeks and Crothers, 1991). The fourth and fifth domains are important for Tat nuclear localization (Ruben et al., 1989), the sixth domain binds to DNA PK and also contribute to viral infectivity (Smith et al., 2003). Importantly in regard to viral latency the functional domains II and III, spanning amino acid positions 22 to 48, are shown to be responsible for transactivation activity (**Figure 2**). The several mutations at positions 22 to 40 amino acid residues (including highly conserved cysteine residues) have been shown to be deleterious with respect to Tat transactivation activity; whereas those at positions 1 to 21 amino acid residues are relatively functionally tolerated (Kuppuswamy et al., 1989; Ruben et al., 1989). Tat plays active role in productive viral replication mainly through enhancement of transcription at viral LTR promoter. Mutational studies have shown there is a strong correlation between Tat transactivation activity and viral replication capacity, whereby the functionally defective Tat has ability to severely inhibit viral replication *in vitro* (Verhoef et al., 1997; Das et al., 2011). This suggests that provirus with functionally defective Tat influences the viral replication and size of the latent reservoir *in vivo*. In respect to the

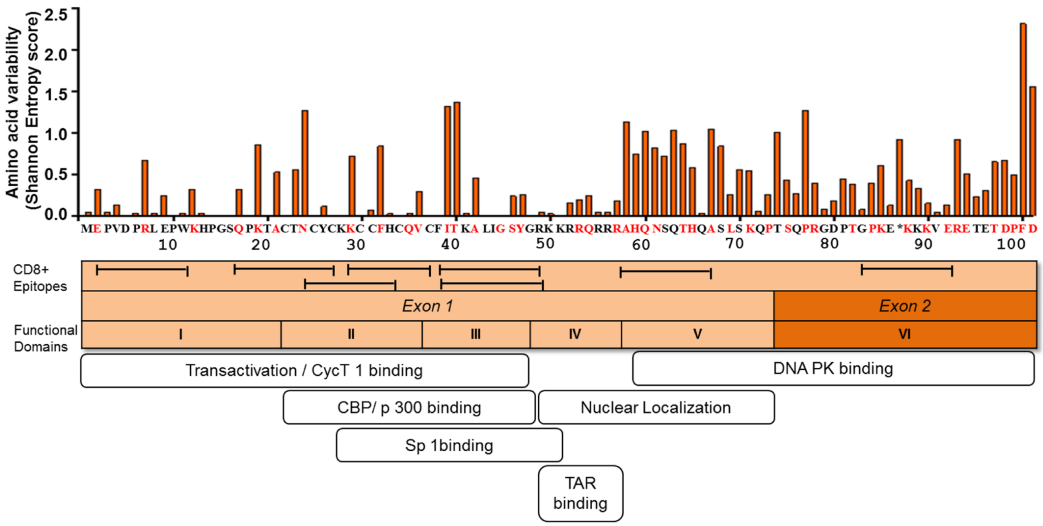


FIGURE 2 | Amino acid variability, immunogenic sites for CTL, and functionally important sites for transactivation activity of HIV-1 Tat. The figure depicts the amino acid variability of Tat as measured by the Shannon entropy score for subtype B Tat sequences ($n = 378$) obtained from a public database (Los Alamos Sequence Database). The immune reactive sites are shown with location of CTL epitopes and the amino acid residues that are associated with CTL escape as determined by statistical association with host HLA class I allele (red). Tat_{HXB2} is used as reference. The domains associated with the transactivation activity are also indicated.

naturally occurring mutations from HIV-1-infected individuals, the Cys-22 to Ser mutation (C22S) in HIV-1 Oyi strain resulted in loss of transactivation activity and was enriched in long-term non-progressive patients (Huet et al., 1989; Peloponese et al., 1999; Watkins et al., 2006). Moreover, several naturally occurring polymorphisms, including P10S, W11R, K19R, A42V, and Y47H, that were observed in 5 HIV-infected subjects at acute or early infection stage, demonstrated impaired transactivation activity and were statistically significantly enriched in the latently infected $CD4^+$ T cells (Yukl et al., 2009). These findings suggest that certain naturally occurring mutations can influence Tat transactivation activity and the establishment of viral latency or reactivation of latent reservoirs during the course of HIV-1 infection *in vivo*. Therefore, this issue warrants for more comprehensive study using a large number of HIV-infected subjects.

Genetic Variability of Tat Driven by Immune-Mediated Selection Forces

It is becoming evident that mutational escape from $CD8^+$ cytotoxic T lymphocyte (CTL) responses represents a potent ongoing driver of global HIV-1 diversification (Price et al., 1997; Goulder et al., 2001; Brumme et al., 2009; Carlson et al., 2012). Tat has also been shown to be frequently targeted by the host HLA-restricted CTL responses (Addo et al., 2001, 2002; Westrop et al., 2009). A number of CTL epitopes have been identified, including PW9 (^{3}PVDPRLEPW^{11}) and EW10 (^{2}EPVDPNLEPW^{11}) restricted by the protective HLA-I alleles, HLA-B*57 and HLA-B*5801, respectively (Schellens et al., 2008; Zhai et al., 2008; Chopera et al., 2011). Additional epitopes are well summarized at the web site, http://www.hiv.lanl.gov/content/immunology/maps/ctl/Tat.html. CTL epitopes are distributed in both highly conserved and polymorphic regions in Tat; however, more number of CTL epitopes are reported at the relatively conserved regions to date (**Figure 2**). A number of Tat mutations in both conserved and variable regions have been reported to be associated with host cellular immune responses in various viral subtypes and host populations (**Figure 2**) (Allen et al., 2000; Guillon et al., 2006; Mason et al., 2009; John et al., 2010; Carlson et al., 2012). Importantly, some of the CTL escape mutations in Tat such as F32L and V36S observed in a frequently recognized (or immunodominant) Tat epitope, CC8 (^{30}CCFHCQVC^{37}) restricted by HLA-C*12:03 (Cao et al., 2003; Liu et al., 2007, 2011), are located at sites that are important for transactivation and co-factor binding (**Figure 2**). Some other CTL escape mutations are located at functionally important regions; N24K, N24T, K29R, and K29S in NF9 (^{24}NCYCKRCCF^{32}) epitope restricted by HLA-A*29:02 (Jones et al., 2004), K40T in FY10 (^{38}FQKKGLGISY^{47}) restricted by HLA-B*15:03 (Liu et al., 2013), and R7S, R7K, and E9D in PW9 (^{3}PVDPRLEPW^{11}) restricted by HLA-A*25:01 (Liu et al., 2007). These data suggest that CTL escape mutations in Tat, especially those located at functionally important conserved regions, have a potential to differentially influence Tat activity. However, it remains elusive as to what extent CTL responses to Tat or CTL escape mutations in Tat may influence viral latency kinetics both at establishment and reversal stages. Also, it is intriguing to ask whether Tat mutations may influence immune recognition of latently infected cells after reactivation. It is also worth to mention that despite the predominant effect of CTL selection pressure on Tat sequence polymorphism, other host immune responses such as those mediated by $CD4^+$ T cells (Lichterfeld et al., 2012; Ranasinghe et al., 2013) and B cells (Goldstein et al., 2001; Moreau et al., 2004) also target Tat; and may therefore

potentially impose selection pressure leading to escape mutations which may differentially affect Tat activity.

CONCLUSION AND FUTURE PERSPECTIVES

To date, the highly genetic viral variability and the existence of latently infected resting $CD4^+$ T lymphocytes and other cells *in vivo* are among the setbacks toward achievement of complete HIV control and eradication. It is generally thought that virus can acquire mutations and evade host immune responses while maintain their fitness effects as minimal as possible. However, similar to the cases in the other HIV-1 proteins such as Gag (Goulder et al., 2001; Troyer et al., 2009) and Nef (Mwimanzi et al., 2013; Kuang et al., 2014), certain naturally occurring immune-associated mutations in Tat may impose fitness cost to the virus. However, it remains poorly described how immune-mediated Tat polymorphisms affect either establishment of viral latency or reactivation of the latently infected cells and also the consequence of such viral polymorphisms on immune recognition. These points could open a new venue to modulate HIV latency and reversal of latency *in vivo* for future therapeutic application toward cure.

AUTHOR CONTRIBUTIONS

DK and TU conceived, designed, compiled the data, and wrote the manuscript.

FUNDING

This work was supported in part by a grant from JSPS KAKENHI Grant number JP16K15284 and JP16H05822, AIDS International Collaborative Research Grant from the Ministry of Education, Science, Sports, and Culture (MEXT) of Japan, and Japan Agency for Medical Research and Development, AMED (Research Program on HIV/AIDS). DK is supported by the scholarship for The International Priority Graduate Programs; Advanced Graduate Courses for International Students (Doctoral Course), MEXT, Japan.

ACKNOWLEDGMENT

The authors also wish to thank M. Mahiti and other lab members for helpful discussion.

REFERENCES

Abbas, W., Tariq, M., Iqbal, M., Kumar, A., and Herbein, G. (2015). Eradication of HIV-1 from the macrophage reservoir: an uncertain goal? *Viruses* 7, 1578–1598. doi: 10.3390/v7041578

Addo, M. M., Altfeld, M., Rosenberg, E. S., Eldridge, R. L., Philips, M. N., Habeeb, K., et al. (2001). The HIV-1 regulatory proteins Tat and Rev are frequently targeted by cytotoxic T lymphocytes derived from HIV-1-infected individuals. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1781–1786. doi: 10.1073/pnas.98.4.1781

Addo, M. M., Yu, X. G., Rosenberg, E. S., Walker, B. D., and Altfeld, M. (2002). Cytotoxic T-lymphocyte (CTL) responses directed against regulatory and accessory proteins in HIV-1 infection. *DNA Cell Biol.* 21, 671–678. doi: 10.1089/104454902760330219

Allen, T. M., O'Connor, D. H., Jing, P., Dzuris, J. L., Mothe, B. R., Vogel, T. U., et al. (2000). Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* 407, 386–390. doi: 10.1038/35030124

Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N., and Margolis, D. M. (2014). Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat. Rev. Microbiol.* 12, 750–764. doi: 10.1038/nrmicro3352

Battistini, A., and Sgarbanti, M. (2014). HIV-1 latency: an update of molecular mechanisms and therapeutic strategies. *Viruses* 6, 1715–1758. doi: 10.3390/v6041715

Brumme, Z. L., John, M., Carlson, J. M., Brumme, C. J., Chan, D., Brockman, M. A., et al. (2009). HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS ONE* 4:e6687. doi: 10.1371/journal.pone.0006687

Canki, M., Thai, J. N., Chao, W., Ghorpade, A., Potash, M. J., and Volsky, D. J. (2001). Highly productive infection with pseudotyped human immunodeficiency virus type 1 (HIV-1) indicates no intracellular restrictions to HIV-1 replication in primary human astrocytes. *J. Virol.* 75, 7925–7933. doi: 10.1128/JVI.75.17.7925-7933.2001

Cao, J., McNevin, J., Malhotra, U., and McElrath, M. J. (2003). Evolution of $CD8^+$ T cell immunity and viral escape following acute HIV-1 infection. *J. Immunol.* 171, 3837–3846. doi: 10.4049/jimmunol.171.7.3837

Carlson, J. M., Brumme, C. J., Martin, E., Listgarten, J., Brockman, M. A., Le, A. Q., et al. (2012). Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J. Virol.* 86, 13202–13216. doi: 10.1128/JVI.01998-12

Cary, D. C., Fujinaga, K., and Peterlin, B. M. (2016). Molecular mechanisms of HIV latency. *J. Clin. Invest.* 126, 448–454. doi: 10.1172/JCI80565

Chakrabarti, L., Hurtrel, M., Maire, M. A., Vazeux, R., Dormont, D., Montagnier, L., et al. (1991). Early viral replication in the brain of SIV-infected rhesus monkeys. *Am. J. Pathol.* 139, 1273–1280.

Chopera, D. R., Mlotshwa, M., Woodman, Z., Mlisana, K., de Assis Rosa, D., Martin, D. P., et al. (2011). Virological and immunological factors associated with HIV-1 differential disease progression in HLA-B 58:01-positive individuals. *J. Virol.* 85, 7070–7080. doi: 10.1128/JVI.02543-10

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., et al. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* 387, 183–188. doi: 10.1038/387183a0

Chun, T. W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D., and Siliciano, R. F. (1995). In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nat. Med.* 1, 1284–1290. doi: 10.1038/nm1295-1284

Coiras, M., Lopez-Huertas, M. R., Perez-Olmeda, M., and Alcami, J. (2009). Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat. Rev. Microbiol.* 7, 798–812. doi: 10.1038/nrmicro2223

Coiras, M., Lopez-Huertas, M. R., Sanchez del Cojo, M., Mateos, E., and Alcami, J. (2010). Dual role of host cell factors in HIV-1 replication: restriction and enhancement of the viral cycle. *AIDS Rev.* 12, 103–112.

Dahabieh, M. S., Battivelli, E., and Verdin, E. (2015). Understanding HIV latency: the road to an HIV cure. *Annu. Rev. Med.* 66, 407–421. doi: 10.1146/annurev-med-092112-152941

Dahmus, M. E. (1996). Phosphorylation of mammalian RNA polymerase II. *Methods Enzymol.* 273, 185–193. doi: 10.1016/S0076-6879(96)73019-7

Das, A. T., Harwig, A., and Berkhout, B. (2011). The HIV-1 Tat protein has a versatile role in activating viral transcription. *J. Virol.* 85, 9506–9516. doi: 10.1128/JVI.00650-11

Davis, L. E., Hjelle, B. L., Miller, V. E., Palmer, D. L., Llewellyn, A. L., Merlin, T. L., et al. (1992). Early viral brain invasion in iatrogenic human immunodeficiency virus infection. *Neurology* 42, 1736–1739. doi: 10.1212/WNL.42.9.1736

Deng, K., Pertea, M., Rongvaux, A., Wang, L., Durand, C. M., Ghiaur, G., et al. (2015). Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* 517, 381–385. doi: 10.1038/nature14053

Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., et al. (1989). Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6925–6929. doi: 10.1073/pnas.86.18.6925

Donahue, D. A., Bastarache, S. M., Sloan, R. D., and Wainberg, M. A. (2013). Latent HIV-1 can be reactivated by cellular superinfection in a Tat-dependent manner, which can lead to the emergence of multidrug-resistant recombinant viruses. *J. Virol.* 87, 9620–9632. doi: 10.1128/JVI.01165-13

Donahue, D. A., Kuhl, B. D., Sloan, R. D., and Wainberg, M. A. (2012). The viral protein Tat can inhibit the establishment of HIV-1 latency. *J. Virol.* 86, 3253–3263. doi: 10.1128/JVI.06648-11

Donahue, D. A., and Wainberg, M. A. (2013). Cellular and molecular mechanisms involved in the establishment of HIV-1 latency. *Retrovirology* 10:11. doi: 10.1186/1742-4690-10-11

Efthymiadis, A., Briggs, L. J., and Jans, D. A. (1998). The HIV-1 Tat nuclear localization sequence confers novel nuclear import properties. *J. Biol. Chem.* 273, 1623–1628. doi: 10.1074/jbc.273.3.1623

Feinberg, M. B., Baltimore, D., and Frankel, A. D. (1991). The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proc. Natl. Acad. Sci. U.S.A.* 88, 4045–4049. doi: 10.1073/pnas.88.9.4045

Feng, S., and Holland, E. C. (1988). HIV-1 tat trans-activation requires the loop sequence within tar. *Nature* 334, 165–167. doi: 10.1038/334165a0

Garber, M. E., Wei, P., KewalRamani, V. N., Mayall, T. P., Herrmann, C. H., Rice, A. P., et al. (1998). The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev.* 12, 3512–3527. doi: 10.1101/gad.12.22.3512

Geng, G., Liu, B., Chen, C., Wu, K., Liu, J., Zhang, Y., et al. (2016). Development of an attenuated tat protein as a highly-effective agent to specifically activate HIV-1 latency. *Mol. Ther.* 24, 1528–1537. doi: 10.1038/mt.2016.117

Goldstein, G., Tribbick, G., and Manson, K. (2001). Two B cell epitopes of HIV-1 Tat protein have limited antigenic polymorphism in geographically diverse HIV-1 strains. *Vaccine* 19, 1738–1746. doi: 10.1016/S0264-410X(00)00393-5

Goulder, P. J., Brander, C., Tang, Y., Tremblay, C., Colbert, R. A., Addo, M. M., et al. (2001). Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 412, 334–338. doi: 10.1038/35085576

Guillon, C., Stankovic, K., Ataman-Onal, Y., Biron, F., and Verrier, B. (2006). Evidence for CTL-mediated selection of Tat and Rev mutants after the onset of the asymptomatic period during HIV type 1 infection. *AIDS Res. Hum. Retroviruses* 22, 1283–1292. doi: 10.1089/aid.2006.22.1283

Huet, T., Dazza, M. C., Brun-Vezinet, F., Roelants, G. E., and Wain-Hobson, S. (1989). A highly defective HIV-1 strain isolated from a healthy Gabonese individual presenting an atypical western blot. *AIDS* 3, 707–715. doi: 10.1097/00002030-198911000-00004

Jeang, K. T., Chun, R., Lin, N. H., Gatignol, A., Glabe, C. G., and Fan, H. (1993). In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor. *J. Virol.* 67, 6224–6233.

John, M., Heckerman, D., James, I., Park, L. P., Carlson, J. M., Chopra, A., et al. (2010). Adaptive interactions between HLA and HIV-1: highly divergent selection imposed by HLA class I molecules with common supertype motifs. *J. Immunol.* 184, 4368–4377. doi: 10.4049/jimmunol.0903745

Jones, N. A., Wei, X., Flower, D. R., Wong, M., Michor, F., Saag, M. S., et al. (2004). Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. *J. Exp. Med.* 200, 1243–1256. doi: 10.1084/jem.20040511

Kuang, X. T., Li, X., Anmole, G., Mwimanzi, P., Shahid, A., Le, A. Q., et al. (2014). Impaired Nef function is associated with early control of HIV-1 viremia. *J. Virol.* 88, 10200–10213. doi: 10.1128/JVI.01334-14

Kumar, A., Abbas, W., and Herbein, G. (2014). HIV-1 latency in monocytes/macrophages. *Viruses* 6, 1837–1860. doi: 10.3390/v6041837

Kuppuswamy, M., Subramanian, T., Srinivasan, A., and Chinnadurai, G. (1989). Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nucleic Acids Res.* 17, 3551–3561. doi: 10.1093/nar/17.9.3551

Lassen, K. G., Ramyar, K. X., Bailey, J. R., Zhou, Y., and Siliciano, R. F. (2006). Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog.* 2:e68. doi: 10.1371/journal.ppat.0020068

Li, G., Piampongsant, S., Faria, N. R., Voet, A., Pineda-Pena, A. C., Khouri, R., et al. (2015). An integrated map of HIV genome-wide variation from a population perspective. *Retrovirology* 12:18. doi: 10.1186/s12977-015-0148-6

Li, J., Chen, C., Ma, X., Geng, G., Liu, B., Zhang, Y., et al. (2016). Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nat. Commun.* 7:11730. doi: 10.1038/ncomms11730

Lichterfeld, M., Gandhi, R. T., Simmons, R. P., Flynn, T., Sbrolla, A., Yu, X. G., et al. (2012). Induction of strong HIV-1-specific CD4+ T-cell responses using an HIV-1 gp120/NefTat vaccine adjuvanted with AS02A in antiretroviral-treated HIV-1-infected individuals. *J. Acquir. Immune Defic. Syndr.* 59, 1–9. doi: 10.1097/QAI.0b013e3182373b77

Lin, X., Irwin, D., Kanazawa, S., Huang, L., Romeo, J., Yen, T. S., et al. (2003). Transcriptional profiles of latent human immunodeficiency virus in infected individuals: effects of Tat on the host and reservoir. *J. Virol.* 77, 8227–8236. doi: 10.1128/JVI.77.15.8227-8236.2003

Liu, M. K., Hawkins, N., Ritchie, A. J., Ganusov, V. V., Whale, V., Brackenridge, S., et al. (2013). Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *J. Clin. Invest.* 123, 380–393. doi: 10.1172/JCI65330

Liu, Y., McNevin, J., Zhao, H., Tebit, D. M., Troyer, R. M., McSweyn, M., et al. (2007). Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: fitness-balanced escape. *J. Virol.* 81, 12179–12188. doi: 10.1128/JVI.01277-07

Liu, Y., McNevin, J. P., Holte, S., McElrath, M. J., and Mullins, J. I. (2011). Dynamics of viral evolution and CTL responses in HIV-1 infection. *PLoS ONE* 6:e15639. doi: 10.1371/journal.pone.0015639

MacDougall, T. H., Shattock, R. J., Madsen, C., Chain, B. M., and Katz, D. R. (2002). Regulation of primary HIV-1 isolate replication in dendritic cells. *Clin. Exp. Immunol.* 127, 66–71. doi: 10.1046/j.1365-2249.2002.01715.x

Marsden, M. D., and Zack, J. A. (2015). Double trouble: HIV latency and CTL escape. *Cell Host Microbe* 17, 141–142. doi: 10.1016/j.chom.2015.01.008

Marzio, G., Tyagi, M., Gutierrez, M. I., and Giacca, M. (1998). HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13519–13524. doi: 10.1073/pnas.95.23.13519

Mason, R. D., De Rose, R., and Kent, S. J. (2009). Differential patterns of immune escape at Tat-specific cytotoxic T cell epitopes in pigtail macaques. *Virology* 388, 315–323. doi: 10.1016/j.virol.2009.03.020

Moreau, E., Belliard, G., Partidos, C. D., Pradezinsky, F., Le Buanec, H., Muller, S., et al. (2004). Important B-cell epitopes for neutralization of human immunodeficiency virus type 1 Tat in serum samples of humans and different animal species immunized with Tat protein or peptides. *J. Gen. Virol.* 85(Pt 10), 2893–2901. doi: 10.1099/vir.0.80365-0

Mousseau, G., Kessing, C. F., Fromentin, R., Trautmann, L., Chomont, N., and Valente, S. T. (2015). The tat inhibitor Didehydro-Cortistatin A prevents HIV-1 reactivation from latency. *MBio* 6:e00465-15. doi: 10.1128/mBio.00465-15

Mwimanzi, P., Markle, T. J., Martin, E., Ogata, Y., Kuang, X. T., Tokunaga, M., et al. (2013). Attenuation of multiple Nef functions in HIV-1 elite controllers. *Retrovirology* 10:1. doi: 10.1186/1742-4690-10-1

Neuveut, C., and Jeang, K. T. (1996). Recombinant human immunodeficiency virus type 1 genomes with tat unconstrained by overlapping reading frames reveal residues in Tat important for replication in tissue culture. *J. Virol.* 70, 5572–5581.

Pagans, S., Pedal, A., North, B. J., Kaehlcke, K., Marshall, B. L., Dorr, A., et al. (2005). SIRT1 regulates HIV transcription via Tat deacetylation. *PLoS Biol.* 3:e41. doi: 10.1371/journal.pbio.0030041

Parada, C. A., and Roeder, R. G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature* 384, 375–378. doi: 10.1038/384375a0

Pearson, R., Kim, Y. K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M., et al. (2008). Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. *J. Virol.* 82, 12291–12303. doi: 10.1128/JVI.01383-08

Peloponese, J. M. Jr., Collette, Y., Gregoire, C., Bailly, C., Campese, D., Meurs, E. F., et al. (1999). Full peptide synthesis, purification, and characterization of six Tat variants. Differences observed between HIV-1 isolates from Africa and other continents. *J. Biol. Chem.* 274, 11473–11478.

Peterlin, B. M., Brogie, J. E., and Price, D. H. (2012). 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdiscip. Rev. RNA* 3, 92–103. doi: 10.1002/wrna.106

Price, D. A., Goulder, P. J., Klenerman, P., Sewell, A. K., Easterbrook, P. J., Troop, M., et al. (1997). Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc. Natl. Acad. Sci. U.S.A.* 94, 1890–1895. doi: 10.1073/pnas.94.5.1890

Rana, T. M., and Jeang, K. T. (1999). Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.* 365, 175–185. doi: 10.1006/abbi.1999.1206

Ranasinghe, S., Cutler, S., Davis, I., Lu, R., Soghoian, D. Z., Qi, Y., et al. (2013). Association of HLA-DRB1-restricted CD4(+) T cell responses with HIV immune control. *Nat. Med.* 19, 930–933. doi: 10.1038/nm.3229

Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M., and Weinberger, L. S. (2015). A hardwired HIV latency program. *Cell* 160, 990–1001. doi: 10.1016/j.cell.2015.02.009

Rossenkhan, R., Novitsky, V., Sebunya, T. K., Musonda, R., Gashe, B. A., and Essex, M. (2012). Viral diversity and diversification of major non-structural genes vif, vpr, vpu, tat exon 1 and rev exon 1 during primary HIV-1 subtype C infection. *PLoS ONE* 7:e35491. doi: 10.1371/journal.pone.0035491

Roy, C. N., Khandaker, I., Furuse, Y., and Oshitani, H. (2015a). Molecular characterization of full-length Tat in HIV-1 subtypes B and C. *Bioinformation* 11, 151–160. doi: 10.6026/97320630011151

Roy, C. N., Khandaker, I., and Oshitani, H. (2015b). Intersubtype genetic variation of HIV-1 Tat Exon 1. *AIDS Res. Hum. Retroviruses* 31, 641–648. doi: 10.1089/AID.2014.0346

Roy, S., Delling, U., Chen, C. H., Rosen, C. A., and Sonenberg, N. (1990). A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev.* 4, 1365–1373. doi: 10.1101/gad.4.8.1365

Ruben, S., Perkins, A., Purcell, R., Joung, K., Sia, R., Burghoff, R., et al. (1989). Structural and functional characterization of human immunodeficiency virus tat protein. *J. Virol.* 63, 1–8.

Ruelas, D. S., and Greene, W. C. (2013). An integrated overview of HIV-1 latency. *Cell* 155, 519–529. doi: 10.1016/j.cell.2013.09.044

Rusnati, M., Tulipano, G., Spillmann, D., Tanghetti, E., Oreste, P., Zoppetti, G., et al. (1999). Multiple interactions of HIV-I Tat protein with size-defined heparin oligosaccharides. *J. Biol. Chem.* 274, 28198–28205. doi: 10.1074/jbc.274.40.28198

Schellens, I. M., Kesmir, C., Miedema, F., van Baarle, D., and Borghans, J. A. (2008). An unanticipated lack of consensus cytotoxic T lymphocyte epitopes in HIV-1 databases: the contribution of prediction programs. *AIDS* 22, 33–37. doi: 10.1097/QAD.0b013e3282f15622

Smith, S. M., Pentlicky, S., Klase, Z., Singh, M., Neuveut, C., Lu, C. Y., et al. (2003). An in vivo replication-important function in the second coding exon of Tat is constrained against mutation despite cytotoxic T lymphocyte selection. *J. Biol. Chem.* 278, 44816–44825. doi: 10.1074/jbc.M307546200

Tong-Starksen, S. E., Baur, A., Lu, X. B., Peck, E., and Peterlin, B. M. (1993). Second exon of Tat of HIV-2 is required for optimal trans-activation of HIV-1 and HIV-2 LTRs. *Virology* 195, 826–830. doi: 10.1006/viro.1993.1438

Troyer, R. M., McNevin, J., Liu, Y., Zhang, S. C., Krizan, R. W., Abraha, A., et al. (2009). Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog.* 5:e1000365. doi: 10.1371/journal.ppat.1000365

Valentin, A., Rosati, M., Patenaude, D. J., Hatzakis, A., Kostrikis, L. G., Lazanas, M., et al. (2002). Persistent HIV-1 infection of natural killer cells in patients receiving highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. U.S.A.* 99, 7015–7020. doi: 10.1073/pnas.102672999

Verhoef, K., Koper, M., and Berkhout, B. (1997). Determination of the minimal amount of Tat activity required for human immunodeficiency virus type 1 replication. *Virology* 237, 228–236. doi: 10.1006/viro.1997.8786

Watkins, J. D., Lancelot, S., Campbell, G. R., Esquieu, D., de Mareuil, J., Opi, S., et al. (2006). Reservoir cells no longer detectable after a heterologous SHIV challenge with the synthetic HIV-1 Tat Oyi vaccine. *Retrovirology* 3:8. doi: 10.1186/1742-4690-3-8

Weeks, K. M., and Crothers, D. M. (1991). RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* 66, 577–588. doi: 10.1016/0092-8674(81)90020-9

Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., and Jones, K. A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* 92, 451–462. doi: 10.1016/S0092-8674(00)80939-3

Westrop, S. J., Qazi, N. A., Pido-Lopez, J., Nelson, M. R., Gazzard, B., Gotch, F. M., et al. (2009). Transient nature of long-term nonprogression and broad virus-specific proliferative T-cell responses with sustained thymic output in HIV-1 controllers. *PLoS ONE* 4:e5474. doi: 10.1371/journal.pone.0005474

Yedavalli, V. S., Benkirane, M., and Jeang, K. T. (2003). Tat and trans-activation-responsive (TAR) RNA-independent induction of HIV-1 long terminal repeat by human and murine cyclin T1 requires Sp1. *J. Biol. Chem.* 278, 6404–6410. doi: 10.1074/jbc.M209162200

Yukl, S., Pillai, S., Li, P., Chang, K., Pasutti, W., Ahlgren, C., et al. (2009). Latently-infected CD4+ T cells are enriched for HIV-1 Tat variants with impaired transactivation activity. *Virology* 387, 98–108. doi: 10.1016/j.virol.2009.01.013

Yusim, K., Kesmir, C., Gaschen, B., Addo, M. M., Altfeld, M., Brunak, S., et al. (2002). Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. *J. Virol.* 76, 8757–8768. doi: 10.1128/JVI.76.17.8757-8768.2002

Zhai, S., Zhuang, Y., Song, Y., Li, S., Huang, D., Kang, W., et al. (2008). HIV-1-specific cytotoxic T lymphocyte (CTL) responses against immunodominant optimal epitopes slow the progression of AIDS in China. *Curr. HIV Res.* 6, 335–350. doi: 10.2174/157016208785132473

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kamori and Ueno. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 April 2017
doi: 10.3389/fmicb.2017.00706

Impact of Genetic Variations in HIV-1 Tat on LTR-Mediated Transcription via TAR RNA Interaction

Larance Ronsard[1,2†], Nilanjana Ganguli[1‡], Vivek K. Singh[3‡], Kumaravel Mohankumar[4,5§], Tripti Rai[6§], Subhashree Sridharan[4,7], Sankar Pajaniradje[4], Binod Kumar[8], Devesh Rai[9], Suhnrita Chaudhuri[10], Mohane S. Coumar[3], Vishnampettai G. Ramachandran[2] and Akhil C. Banerjea[1*]*

[1] Laboratory of Virology, National Institute of Immunology, Delhi, India, [2] Department of Microbiology, University College of Medical Sciences and Guru Teg Bahadur Hospital, Delhi, India, [3] Centre for Bioinformatics, School of Life Sciences, Pondicherry University, Pondicherry, India, [4] Department of Biochemistry and Molecular Biology, Pondicherry University, Pondicherry, India, [5] Department of Veterinary Physiology and Pharmacology, Texas A&M University, College Station, TX, USA, [6] Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences, Delhi, India, [7] Department of Symptom Research, The University of Texas MD Anderson Cancer Center, Houston, TX, USA, [8] Department of Microbiology and Immunology, Rosalind Franklin University of Medicine and Science, Chicago, IL, USA, [9] Department of Microbiology, All India Institute of Medical Sciences, Delhi, India, [10] Department of Neurological Surgery, Northwestern University, Chicago, IL, USA

HIV-1 evades host defense through mutations and recombination events, generating numerous variants in an infected patient. These variants with an undiminished virulence can multiply rapidly in order to progress to AIDS. One of the targets to intervene in HIV-1 replication is the *trans*-activator of transcription (Tat), a major regulatory protein that transactivates the long terminal repeat promoter through its interaction with *trans*-activation response (TAR) RNA. In this study, HIV-1 infected patients ($n = 120$) from North India revealed Ser46Phe (20%) and Ser61Arg (2%) mutations in the Tat variants with a strong interaction toward TAR leading to enhanced transactivation activities. Molecular dynamics simulation data verified that the variants with this mutation had a higher binding affinity for TAR than both the wild-type Tat and other variants that lacked Ser46Phe and Ser61Arg. Other mutations in Tat conferred varying affinities for TAR interaction leading to differential transactivation abilities. This is the first report from North India with a clinical validation of CD4 counts to demonstrate the influence of Tat genetic variations affecting the stability of Tat and its interaction with TAR. This study highlights the co-evolution pattern of Tat and predominant nucleotides for Tat activity, facilitating the identification of genetic determinants for the attenuation of viral gene expression.

Keywords: HIV-1 Tat, transactivation, TAR RNA, genetic variations, recombination, mutations

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Hirotaka Ode,
Nagoya Medical Center (NHO), Japan
Aurelio Cafaro,
Istituto Superiore di Sanità, Italy
Takao Masuda,
Tokyo Medical and Dental University, Japan

***Correspondence:**
Akhil C. Banerjea
akhil@nii.ac.in;
akhil@nii.res.in
Larance Ronsard
laraphds@gmail.com;
LRonsard@mgh.harvard.edu

†Present address:
Larance Ronsard,
Ragon Institute of MGH, MIT and Harvard, 400 Technology Square, Cambridge, MA, USA

‡Nilanjana Ganguli and Vivek K. Singh have joint 2nd authorship.

§Kumaravel Mohankumar and Tripti Rai have joint 3rd authorship.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *07 February 2017*
Accepted: *05 April 2017*
Published: *21 April 2017*

Citation:
Ronsard L, Ganguli N, Singh VK, Mohankumar K, Rai T, Sridharan S, Pajaniradje S, Kumar B, Rai D, Chaudhuri S, Coumar MS, Ramachandran VG and Banerjea AC (2017) Impact of Genetic Variations in HIV-1 Tat on LTR-Mediated Transcription via TAR RNA Interaction. Front. Microbiol. 8:706. doi: 10.3389/fmicb.2017.00706

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) overcomes the host immune defense by rapid evolution and genetic diversification (Dougherty and Temin, 1988; Jetzt et al., 2000). Rapid replication without the benefit of proof-reading leads to the generation of a large number of mutations and recombination events (Wolinsky et al., 1996; Jetzt et al., 2000). This vigorous recombination allows HIV-1 to produce multiple groups, subtypes, sub-subtypes and recombinants in an infected patient (Wolinsky et al., 1996; Buonaguro et al., 2007). Of these, only highly virulent variants and recombinants are likely to adapt and spread in a population

(Ho et al., 1995; Buonaguro et al., 2007). It is known that genetic and functional changes in HIV-1 genes accomplished with an emergence of recombination events (Blackard et al., 2002), could affect the virulence of HIV-1 resulting in vulnerability to AIDS despite antiretroviral therapy (ART) (Geretti, 2006; Kirchhoff, 2009; Sharp and Hahn, 2011; Santoro and Perno, 2013).

Among HIV-1 proteins, *trans*-activator of transcription (Tat) plays an important role in mediating the viral transcription (Okamoto et al., 1996; Zhu et al., 1997). Tat is expressed during the early stages of infection, which is encoded by two exons (exons 1 and 2) in a multiple spliced mRNA (Arya et al., 1985). Tat strongly activates transcription (Muesing et al., 1987) from long terminal repeat (LTR) promoter through a strong interaction with *trans*-activation response element (TAR) sequence at the 5′ end of the LTR (+1 to +59) (Dingwall et al., 1989; Buonaguro et al., 1994). During the binding to TAR and host factors Cdk9 and cyclin T1, Tat alters the transcription complex, recruiting a positive transcription elongation complex (P-TEFb), an elongation factor composed of cyclin T1 (CycT1) and Cdk9 that phosphorylates the C-terminal domain of RNA polymerase II leading to the increased production of viral RNA (Zhou and Rana, 2002). Tat contributes to the pathogenesis of HIV-1 through its pivotal role in replication, T-cell apoptosis, co-receptor regulation, cytokine induction, and other viral activities in the host cells (Fulcher and Jans, 2003; Romani et al., 2010).

HIV-1 produces highly divergent variants within a single patient, thereby overcoming the host immune responses, antiretroviral restriction factors, and other selection mechanisms (Konings et al., 2006; Coffin and Swanstrom, 2013). In India, the dominant HIV-1 type is subtype C (>95%), with the emergence of recombinants like A/C, A/E, and B/C (<2%) (Neogi et al., 2011; Ronsard et al., 2015), indicating that HIV-1 is under the stress toward the positive selection in the Indian population. Therefore, it is important to understand the nature of HIV-1 evolution in the population. Previous data showed that the genetic variations in Tat could lead to varying levels of LTR-driven transcription (Ronsard et al., 2014); however, how Tat variants differ in their transactivation abilities have not been explored. Tat is also known for its interaction with various cellular proteins and for its effect on modulating the viral gene expression, which in turn enhances the virulence (Bucci, 2015; Mediouni et al., 2015; Yuan et al., 2015) indicating one of the suitable targets against HIV-1 infection (Hamy et al., 2000; Burton et al., 2004; Nunnari et al., 2008; Sun et al., 2012; Ensoli et al., 2014). We earlier reported that Tat is a highly conserved protein with only a few genetic variations in the functional domains. However, a novel Ser46Phe mutation is prevalent among North Indian population (Ronsard et al., 2014). More than 20% of HIV-1 infected patients carry this mutation along with other mutations; therefore it is important to understand the role of these variants from the population.

Here, we report the differential ability of Tat variants with Ser46Phe mutation to significantly enhanced LTR transactivation ($P < 0.05$) compared to wild-type and other Tat variants that lack this mutation. We observed that variants with this mutation exhibited a strong interaction with TAR by *in vitro* and *in silico* studies, whereas other Tat variants exhibited varying levels of Tat–TAR mediated transactivation. Molecular dynamics (MDs) simulation data confirmed that Ser46Phe mutation exhibits a strong binding with TAR. Here, we illustrate how HIV-1 virus has evolved during selection pressure in the North Indian population with various mutations to adapt and survive in the host cells by enhancing its functional activity.

RESULTS

Selection of Tat Variants for Functional Characterization

HIV-1 specimens from 120 patients were collected, and the Tat gene was amplified by polymerase chain reaction (PCR) and sequenced as described in Section "Materials and Methods." From 120 variants, 15 variants were chosen based on mutations in the Tat gene and these variants were segregated into three groups (selected at least 3 variants in each group from total 120 variants which consisted of similar nucleotide changes) for LTR transactivation study (**Figure 1A**). Next, three Tat variants (TatN12, TatD60, and TatVT6) were selected based on their similarity in inducing LTR transactivation and carrying similar mutations (similar pattern of nucleotide changes) in the Tat gene (selected a variant from each group as a representative variant) for the TAR RNA interaction study which included TatN12, a subtype C variant (that lacked Ser46Phe) with Leu35Pro and Gly44Ser; TatD60, also a subtype C variant (with Ser46Phe) with Glu9Lys and Ser61Arg; and TatVT6, a B/C recombinant (that lacked Ser46Phe) (**Figure 1B**). Three groups were chosen based on the similarities in their genetic and functional activities of Tat (variants with similar nucleotide changes resulting in similar levels of LTR transactivation). It is worth mentioning that Ser46Phe mutation is also reported from neighboring countries like Myanmar and China (HIV database)[1]; however, the functional implication of this mutation on Tat–TAR mediated transactivation has not been studied. The phylogenetic tree was constructed with Tat variants to explain the Tat genetic variations occurring in the North Indian population (Supplementary Figure S1). Notably, the three variants (TatD60, TatE59, and TatE64 were used for transactivation study) and other variants (TatVT1, TatVT3, TatA7, TatN14, TatN17, TatCSW1, TatS5, TatS6) with Ser46Phe were clustered together in the phylogenetic tree indicating the proper classification of Tat variants into three different groups for the functional characterization. Further, the recombination event was confirmed in the Tat variants (TatVT6) using RIP (Recombinant Identification Program) with a confidence threshold 90% and a window size of 100 (Supplementary Figure S2).

Role of Tat Mutations on Viral Transcription

To determine the transactivation activity of Tat variants, a luciferase assay was performed in HEK293 cells with 15 selected

[1]http://www.hiv.lanl.gov/

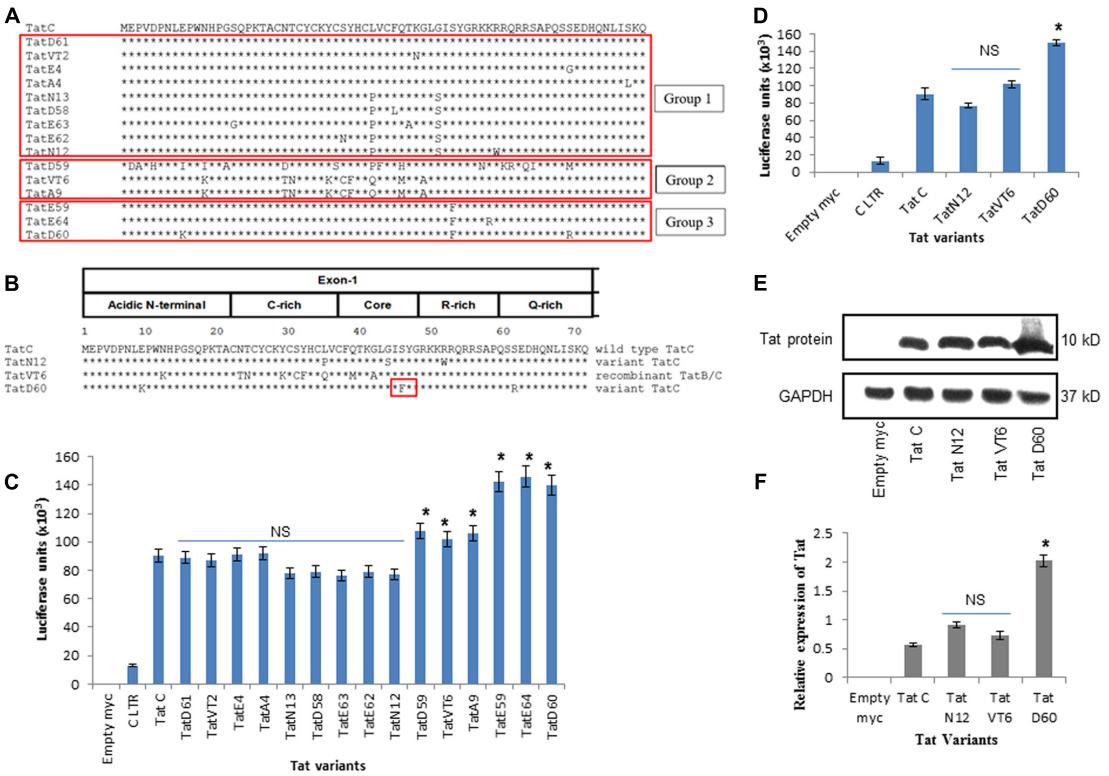


FIGURE 1 | HIV-1 LTR transactivation by Tat variants. (A) Tat variants from HIV-1 infected individuals (n = 15) were aligned with wild-type Tat C (C.IN.93.93IN905). **(B)** Representative data of Tat variants[TatN12 (lack Ser46Phe) with Leu35Pro and Gly44Ser, TatVT6 (lack Ser46Phe) with B/C recombination and TatD60 (with Ser46Phe)] were aligned with wild-type Tat C. **(C)** HEK293 cells were co-transfected with pCMV-myc Tat variants and pGL3-Luc subtype C LTR. After 24 h of transfection, cells were harvested and lysed and luciferase activity was measured. The relative transactivation was expressed as mean luciferase units. Wild-type TatC was used as a reference Tat for comparison. Empty myc vector was used as a control. Subtype C LTR was used as a loading control. Luciferase activity of 15 Tat variants and wild-type Tat C normalized to empty pCMV-myc vector. **(D)** Luciferase activity of unique Tat variants and wild-type TatC normalized to empty pCMV-myc vector. Expression of Tat variants in HEK293 cells. HEK293 cells were transfected with pCMV-myc Tat variants and wild-type TatC. After 24 h, cells were harvested and immunoblotted with anti-myc antibody. The relative Tat protein expression was measured using ImageJ software. Wild-type TatC was used as a reference Tat for comparison. Empty myc vector was used as a control. GAPDH was used as a loading control. **(E)** Relative protein expression of Tat variants and wild-type TatC. **(F)** Quantification of Tat protein expression normalized to GAPDH. Error bar represents the standard deviation in triplicates. Statistical comparison of each Tat variant to Tat C was calculated by one-way ANOVA with the Tukey's test (* denotes $P < 0.05$ and NS denotes not significant).

Tat variants of three groups. Wild-type TatC was used as a reference for baseline transactivation levels of Tat variants. Tat variants with similar nucleotide changes from all three groups (n = 15) showed similar levels of LTR transactivation suggesting that Tat-induced transactivation is dependent on the genetic determinants of Tat variants (**Figure 1C**). We observed that the levels of transactivation induced by TatN12 were similar or slightly lower than the wild-type TatC but not significantly; TatVT6-induced LTR transactivation was a slightly higher than wild-type TatC ($P < 0.05$). TatD60 carrying Ser46Phe showed a significantly higher level of transactivation ($P < 0.05$) than wild-type Tat C as well as other variants. In summary, the order of transactivation induced by variants TatD60 > TatV6 > TatN12 (**Figure 1D**) indicating Ser46Phe mutation play a significant role in transactivation, whereas, other natural mutations did not have a significant role on LTR transactivation. Similar pattern of transactivation was observed with the Tat variants along with the subtype B LTR (Supplementary Figure S3).

Tat Mutations Affect Tat Protein Expression

To examine the effect of Tat variants on intracellular protein expression, Tat variants (TatN12, TatVT6, and TatD60) and wild-type Tat C were cloned into a pCMV-myc vector and were expressed and detected by immunoblotting with anti-myc antibody. Wild-type Tat C was used as a reference for comparison of protein expression. TatN12 and TatVT6 were expressed at similar levels of protein expression when compared to wild-type Tat C; while TatD60 (Ser46Phe) resulted in an elevated level of expression (**Figures 1E,F**) suggesting that Ser46Phe and Ser61Arg mutations in TatD60 could affect Tat protein expression.

Tat Mutations Alter TAR Interaction

To examine the ability of Tat variants to interact with TAR, Tat variants were over-expressed and purified from *Escherichia coli* BL21 strain (Supplementary Figure S4). These proteins

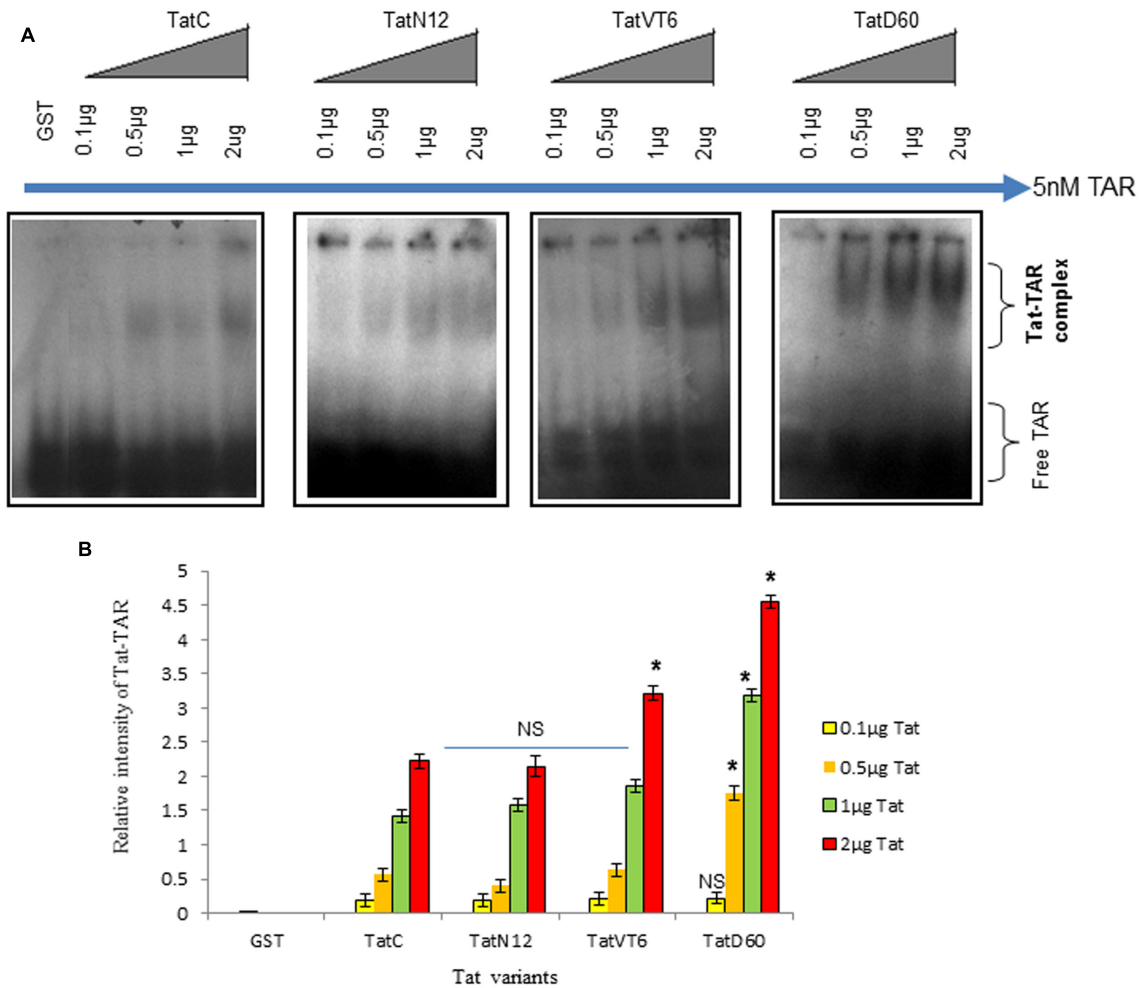


FIGURE 2 | Tat–TAR interaction by EMSA. Tat variants and wild-type Tat cloned in pGEX-4T2 vector were expressed in BL21 (DE3) pLysS cells. Expressed Tat proteins were isolated and verified by immunoblotting with anti-Tat antibody. Increasing concentrations of purified Tat proteins were subjected to interact with subtype B TAR and these Tat ± TAR complexes were detected by autoradiography. The relative intensity of Tat variants–TAR complexes were measured using ImageJ software. Wild-type TatC–TAR complex was used as a reference Tat for comparison. GST was used as a control. Free TAR was used as a loading control. **(A)** Relative intensity of Tat variants–TAR complexes and wild-type TatC–TAR complex. **(B)** Quantification of Tat variants–TAR complexes normalized to empty vector. Error bar represents the standard deviation in triplicates. Statistical comparison of each Tat variant to Tat C was calculated by one-way ANOVA with the Tukey's test (* denotes $P < 0.05$ and NS denotes not significant).

were incubated with TAR (synthesized by *in vitro* transcription) and Tat–TAR binding affinity was measured by Non-denaturing PAGE. Wild-type TatC–TAR complex formation was used as a reference. The binding affinity of TatN12 to TAR was less than wild-type TatC–TAR complex but not significantly; whereas the TatD60 variant carrying Ser46Phe showed significantly higher ($P < 0.05$) binding affinity with TAR RNA (**Figures 2A,B**). The binding affinity of the TatVT6–TAR complex was slightly higher than wild-type TatC–TAR complex but did not reach statistical significance.

Tat Genetic Variations Determine the Protein Stability

To determine whether Ser46Phe and other mutations play a role on Tat protein stability, a cycloheximide chase assay was performed. Briefly, HEK293 cells were transfected with Tat variants and incubated with cycloheximide for different time intervals, and the amount of Tat in cell lysates was quantified by western blotting. Wild-type Tat C protein stability was treated as baseline. After 3 h of cycloheximide treatment, the level of TatN12 was found to be reduced. In contrast, the protein levels of TatD60 and TatVT6 were detectable even after 3 h of treatment (**Figures 3A,B**) indicating that Ser46Phe and recombination events in TatD60 and TatVT6 respectively could possibly stabilize Tat protein.

Influence of Genetic Variability on Tat Ubiquitination

To understand whether ubiquitination determines the differential stability of Tat proteins, we measured the level

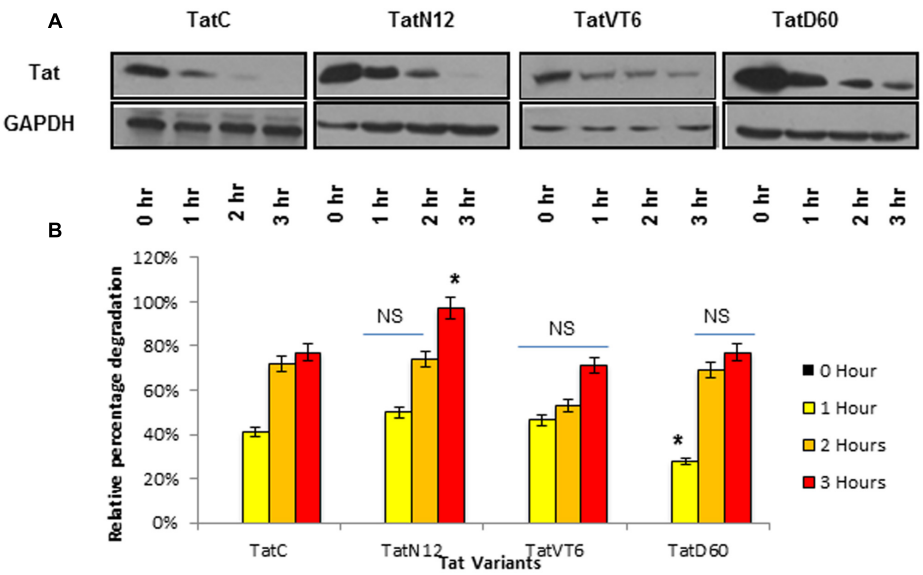


FIGURE 3 | Intracellular stability of Tat variants. HEK293 cells were transfected with pCMV-Myc Tat variants and wild-type TatC, cycloheximide (100 μg/ml) was added after 24 h and cells were harvested at different time intervals and immunoblotted with anti-Tat antibody. The relative percentage of Tat protein degradation was measured using ImageJ software. Wild-type TatC was used as a reference Tat for comparison. Tat proteins expressed after 24 h of transfection (before adding cycloheximide) were used as controls. GAPDH was used as a loading control. **(A)** Relative protein expression of Tat variants and wild-type TatC at different time intervals. **(B)** Quantification of Tat protein degradation normalized to GAPDH. Error bar represents the standard deviation in triplicates. Statistical comparison of each Tat variant to Tat C was calculated by one-way ANOVA with the Tukey's test (* denotes $P < 0.05$ and NS denotes not significant).

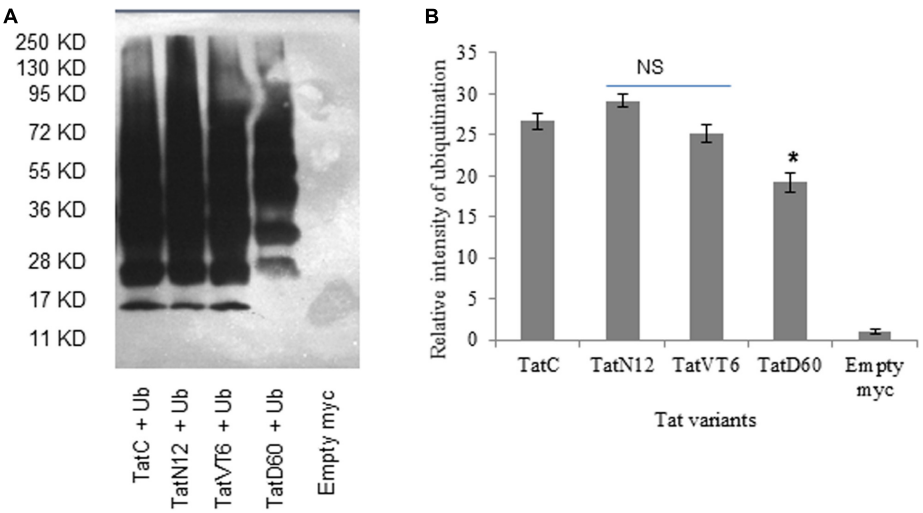


FIGURE 4 | Ubiquitination of Tat variants. HEK293 cells were co-transfected with His6-Ubiquitin protein, and pCMV-myc Tat variants and wild-type TatC. After 24 h, cells were treated with MG132 (25 μM) for 8 h. Cell lysates were incubated with Ni-NTA beads and immunoblotted with anti-myc antibody. The relative intensity of Tat protein ubiquitination was measured using ImageJ software. Wild-type TatC was used as a reference Tat for comparison. Empty vector was used as a control. **(A)** Relative intensity of Tat proteins and wild-type Tat protein ubiquitination. **(B)** Quantification of Tat ubiquitination normalized to empty myc vector. Error bar represents the standard deviation in triplicates. Statistical comparison of each Tat variant to Tat C was calculated by one-way ANOVA with the Tukey's test (* denotes $P < 0.05$ and NS denotes not significant).

of ubiquitination of Tat variants in HEK293 cells. Wild-type Tat C ubiquitination was treated as baseline. TatN12 resulted in slightly higher or similar level of ubiquitination compared to wild-type Tat indicating less stability of this protein than TatD60 and TatVT6 proteins. TatVT6 showed a slightly lower ubiquitination than wild-type TatC while TatD60 showed less ubiquitination indicating that Ser46Phe and Ser61Arg could stabilize Tat protein (**Figures 4A,B**), though it is not known if Ser46 and Ser61 are targets for ubiquitination (Wang et al., 2007). The differential levels of ubiquitination of Tat variants

appear to be dependent on the intracellular level of Tat protein expression.

Tat Mutations Govern Stable Interaction with TAR

Next, we performed MD simulation to estimate the stability of TatD60–TAR and TatC–TAR complexes. We have made homology models of wild-type Tat and TatD60, and then docked with TAR RNA. TatC–TAR complex was treated as a baseline stability of the complex. MD simulation recorded the trajectories of Tat protein backbones and TAR in the docked complex for 20 nanoseconds (ns) in an aqueous environment. Root mean square deviation (RMSD) of Tat protein backbone and TAR ribonucleotides at $t = 0$ ns and $t = 20$ ns were compared to predict the stability of the complexes.

Root mean square deviation of TatD60 backbone was observed at around 5 Å and remained stable throughout the simulation, except for a brief period of higher shift between 8 and 10 ns (**Figure 5A**). In contrast, RMSD for TatC backbone showed more fluctuations. In particular, the upward movement observed after 8 ns was not completely stabilized at the end of the simulation. The average length of protein backbone RMSD for the final 5 ns of simulation, 5.2 Å for TatD60 and 5.7 Å for wild-type TatC, indicated that TatD60–TAR complex had higher stability at equilibrium than wild-type TatC–TAR complex.

Root mean square deviation of TAR ribonucleotides in TatD60–TAR complex showed more drift when compared to TatC–TAR complex (**Figure 5B**), particularly from 8 to 10 ns. The comparison of extracted TAR structure alone at different time frames showed that RMSD was reduced by 19.3 Å in TatD60–TAR complex and by 7.8 Å in wild-type TatC–TAR complex. This reduction could be due to Ser46Phe mutation in TatD60 which created a conformational change in order to bind strongly with TAR ribonucleotides and produced a more stable TatD60–TAR complex.

Next, root mean square fluctuation (RMSF) of each amino acid in Tat proteins differed between TatD60 and TatC. For TatC, maximum fluctuations were observed for five residues at Ser31, Tyr32, His33, Lys41, and Gly42 in addition to the N-terminal and the C-terminal residues (**Figure 5C**). For TatD60, maximum fluctuations were observed for only three residues at Gly15, Ser16, and Lys19 in addition to the N-terminal residues (**Figure 5C**). RMSF was lower for Ser70 and Lys71 in TatD60 when compared to wild-type TatC.

In the case of TAR, ribonucleotides in the bulge region (U23, C24, and U25) involved in Tat interaction, have similar RMSF when interacting with TatC or TatD60, however the loop ribonucleotide G32 appeared to be more flexible in TatC–TAR complex (**Figure 5D**). RMSF of each TAR ribonucleotides in TatD60–TAR complex showed less drift when compared to TatC–TAR complex (**Figure 5D**).

Hydrogen Bonds in Tat–TAR Complexes

In wild-type TatC, Ser46 resulted in a single H-bond with TAR (occupancy ∼71%) (**Figure 5E**), while the corresponding position was Phe46 in TatD60 resulting in one H-bond with TAR only transiently (occupancy below 1%); evidently, the single Ser46Phe mutation had a considerable effect on the H-bond profile of the TatD60–TAR complex (**Figure 5F**). In case of ribonucleotides in TAR, the wild-type Tat had less interaction (**Figure 5G**) than compared to TatD60 (**Figure 5H**) indicating strong H-interaction of TatD60 toward TAR.

Further, analysis of simulated structures at 20 ns revealed potentially critical differences in Tat–TAR interaction at bulge ribonucleotides A22, U23, C24, and U25. TatD60–TAR complex showed stacking interaction between U23 and A22, and standard base-pairing between coplanar bases A22 and U40 (**Figure 5I**) whereas both these features were missing in the TatC–TAR complex (**Figure 5J**). Overall, the number of hydrogen bonds formed between Tat variants and TAR represented in the order of TatD60 > TatVT6 > TatN12.

Binding Free Energy and its Residue Wise Decomposition Analysis of Tat–TAR Interaction

To understand the changes in free energy during formation of Tat variants–TAR, complexes were computed using residue-wise energy decomposition analysis by MM/GBSA method. TatC–TAR was used as a reference for Tat–TAR complex free energy decomposition. ΔG_{bind} of TatD60–TAR was estimated to be −80.49 kcal/mol, whereas in TatC–TAR ΔG_{bind} was −67.14 kcal/mol (Supplementary Table S1). The main forces contributing to the difference between ΔG_{bind} for TatD60–TAR and TatC–TAR were van der Waals and electrostatic interactions.

Residue-wise binding energy decomposition (**Figure 5K** and Supplementary Table S2) showed significant contributions by the following residues in TatD60: Leu35 (−6.447 ± 9.452), Arg49 (−12.070 ± 15.687), Lys51 (−5.266 ± 18.739), Arg53 (−13.688 ± 12.174), Gln54 (−8.058 ± 9.557), and Asn67 (−8.708 ± 6.229), and the following residues in TatC: Arg49 (−16.524 ± 17.477), Lys51 (−7.876 ± 14.652), Arg52 (−10.619 ± 16.843), and Arg53 (−6.192 ± 24.679). In contrast, both TatC and TatD60, Gln54 (0.065 ± 7.828) and Asn67 (0.181 ± 7.948) disfavored the binding ability. These findings further support evidence for a higher affinity of TatD60 for TAR.

Comparison of Crystal Structure and Modeled Structure of Tat–TAR Complex

Further, we have analyzed the Tat crystal structure (PDB ID: 5L1Z) with TAR RNA (**Figure 5L**) and then the complex was compared to the Tat homology modeled complex with TAR RNA (**Figure 5M**). In spite of having few differences in the Tat modeled protein and in the Tat crystal structure, their binding pattern with TAR RNA were found to be similar and also both the complexes were following the same trend in making the hydrogen bonds. The residues Lys50, Lys 51, Arg52, Arg55, Ser57, Asn67, and Ser70 are the major residues contributing more toward the Tat and TAR binding were the same in corresponding residues in the crystal structure (PDB ID: 5L1Z). The additional hydrogen bonds formed in the crystal structure were Thr40 and Tyr46. Those additional hydrogen bonds with hydrogen bond distance 2.5 and 2.8 Å

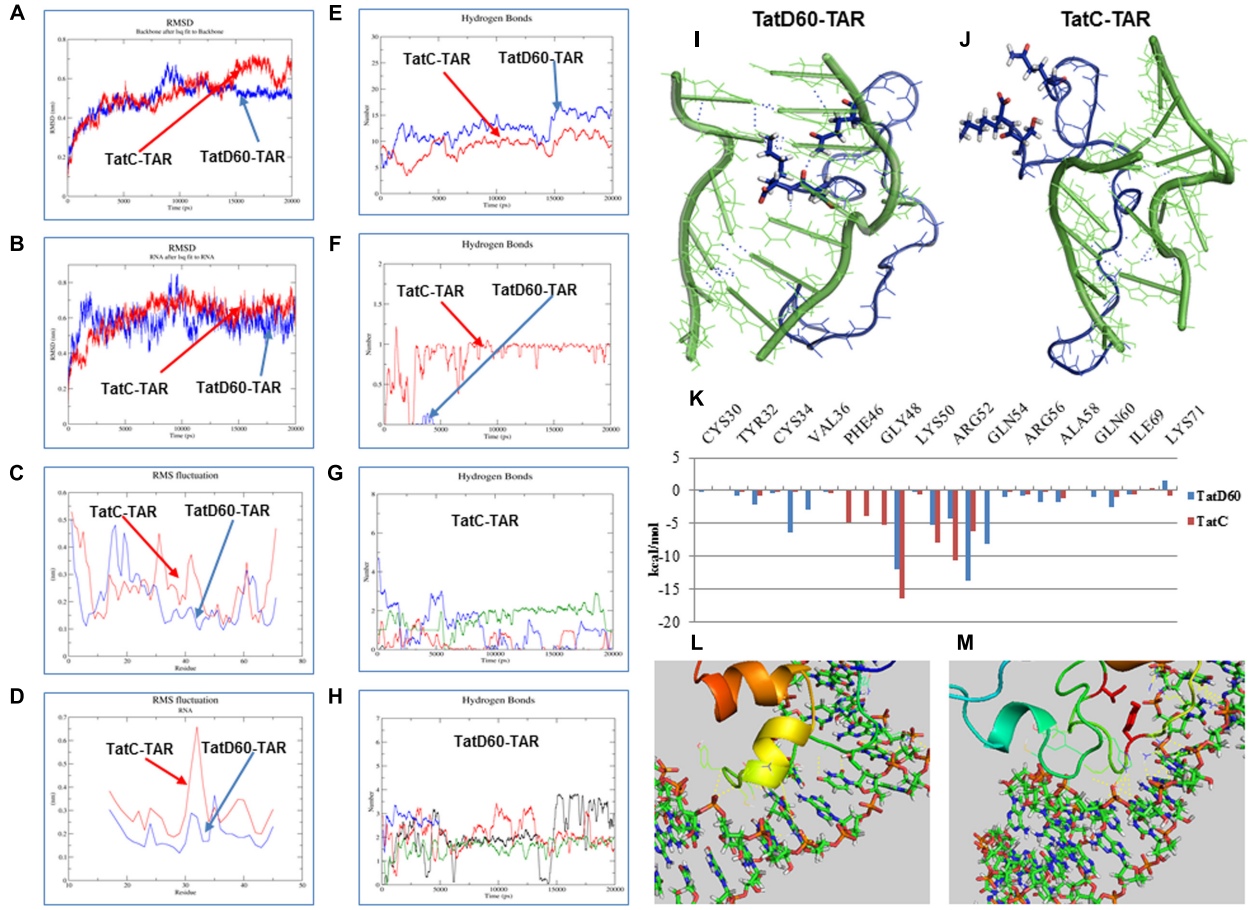


FIGURE 5 | Molecular dynamic (MD) simulation of Tat protein with TAR. MD simulation for Tat–TAR was carried out for 20 ns to predict root mean square deviation (RMSD) of Tat variants. **(A)** RMSD of Tat residues in wild-type TatC–TAR (Red line) and TatD60–TAR (Blue line). **(B)** RMSD of TAR ribonucleotides in wild-type TatC–TAR (Red line) and TatD60–TAR (Blue line). **(C)** Root mean square fluctuation (RMSF) of Tat individual residues in wild-type TatC–TAR (red line) and TatD60–TAR (blue line). **(D)** RMSF of TAR individual ribonucleotides in wild-type TatC–TAR (red line) and TatD60–TAR (blue line). Hydrogen bonds in Tat–TAR complex. Number of H-bonds formed between Tat variants and TAR per time frame was measured; **(E)** number of H-bonds in wild-type TatC–TAR complex (red line) and TatD60–TAR complex (blue line). **(F)** Number of H-bonds formed at Ser46th position in wild-type TatC–TAR complex (red line) and at Phe46 in TatD60–TAR complex (blue line). **(G)** Number of H-bonds formed by TAR ribonucleotides of A22 (black line), U23 (red line), C24 (green line), and U25 (blue line) in wild-type TatC–TAR complex. **(H)** Number of H-bonds formed by TAR ribonucleotides of A22 (black line), U23 (red line), C24 (green line) and U25 (blue line) in TatD60–TAR complex. Tat variants interaction with TAR during simulation. Tat protein (blue) and TAR (green) interaction in the bulge region at the end of 20 ns simulation was captured. H-bonds represented in dotted blue lines. **(I)** Structure of wild-type TatC–TAR complex during 20 ns simulation. **(J)** Structure of TatD60–TAR complex during 20 ns simulation. Binding free energy of Tat–TAR complex. **(K)** Relative residue-wise energy contribution for binding wild-type TatC–TAR and TatD60–TAR complexes. Only selected residues with important contribution are shown. Comparison of crystal structure and modeled structure of Tat–TAR complex. Comparison of docked structure of Tat–TAR complex with homology modeled Tat structure and with the crystal structure (PDB ID: 5L1Z) of Tat. **(L)** Crystal structure complex of TatD60–TAR. **(M)** Homology modeled structure complex of TatD60–TAR.

were distinct from the Tat modeled structure as it was modeled as extended loop and it was away from the C19 and U22 respectively.

Clinical Status of HIV-1 Infected Patients with Tat Variants

In this study, CD4 counts were follow-up for 6 months for HIV-1 infected patients ($n = 15$) which included: group 1 was consisted of HIV-1 infected patients ($n = 9$) lacking Ser46Phe in Tat C; group 2 was consisted of HIV-1 infected patients ($n = 3$) lacking Ser46Phe in Tat B/C recombinant; and group 3 was consisted of HIV-1 infected patients ($n = 3$) with Ser46Phe and Ser61Arg in Tat C.

Of these 15 patients, the mean CD4 count in group 1 patients ($n = 9$) was 470, and after 6 months of ART, the mean value of CD4 count increased to 557 with ±87 standard deviation of the mean value. This showed that the patients in group 1 benefited from ART and indicated that there was no effect of Tat mutations on CD4 counts. However, two patients (TatE4 and TatD58) had almost constant CD4 counts while the other seven patients had increased CD4 counts. Next, the mean CD4 count in group 2 patients ($n = 3$) was 613, and after 6 months of ART, the mean value of CD4 count decreased to 578 with ±35

standard deviation of the mean value. This showed that the CD4 counts of the patients in group 2 could be affected by B/C Tat; however, the changes were not significant. In group 3 patients ($n = 3$), the mean CD4 count was 493. After 6 months of ART, the mean value of CD4 count decreased to 359 with ± 134 standard deviation of the mean value. This showed that the patients in this mutation group 3 could be affected due to Ser46Phe and other mutations in this group (Supplementary Table S3).

In the total of 15 patients, the mean viral load after ART in group 1 patients ($n = 9$) was less than 50 copies/ml. This showed that the patients did not affected by the transactivation level induced by Tat mutations in this group (particularly TatN12 is <50 copies/ml). However, two patients (TatE4 and TatD61) had slightly detectable viral loads but not significantly. Next, the mean viral load in group 2 patients ($n = 3$) was slightly increased to 158 copies/ml. This showed that the viral load of the patients could be affected by the transactivation level induced by B/C Tat in this group (particularly TatVT6 is 182 copies/ml). In group 3 patients ($n = 3$), the mean viral loads was 265 copies/ml indicating that the patients in this mutation group 3 could be affected by the transactivation level caused due to Ser46Phe and other mutations in this group (particularly TatD60 is 346 copies/ml), however, there could be also multiple other factors might affect both the CD4 count and the viral load indicating the need for the further study in a large sample size (Supplementary Table S3).

DISCUSSION

A rate-limiting factor in the management of HIV infections, is the plethora of genetic variations leading to failure of clinical trials (Santoro and Perno, 2013). Each geographical region has its own profile of HIV-1 types, subtypes, recombinants and mutations, and these genetic variations lead to differential potential in inducing HIV-1 pathogenesis (Rodriguez et al., 2009). We previously reported the differential expression of certain viral genes of subtypes B and C with respect to their functions (Sood et al., 2008; Gupta and Banerjea, 2009). The viral or cellular components responsible for these differences have not been studied in details and in most cases, the molecular details of these genetic determinants have not been well-explored. Our previous reports showed that the functional role of viral proteins is mediated by the genetic determinants present in the viral genes (Siu et al., 2013; Zhu et al., 2013). It is known that Tat subtype-specific variations exhibit widely differing viral activities including their ability to activate the HIV-1 LTR promoter (Li et al., 2012); however, the functional and clinical consequences of the substantial genetic variations of Tat occurring in North India have not been studied. The growing number of studies on Tat-based inhibitors against HIV-1 replication (Hamy et al., 1997; Lalonde et al., 2011) indicates the importance of this study.

In this regard, we attempted to understand the differential ability of Tat variants to activate LTR transactivation through their ability to interact with TAR using *in vitro* and *in silica* approaches. Tat protein is highly conserved protein among North Indians. For viral replication, it is important that Tat retains its functional activity because any changes in the genetic makeup of Tat could lead to drastic modulation in transcription-coupled pathogenicity. In our survey of 120 Tat sequences from 120 HIV-1 infected patients, we found certain point mutants and B/C recombinants commonly occurring among North Indians. In order to understand the role of these natural mutations, we carried out genetic and functional analyses in correlation with clinical CD4 counts. Both *in silico* and *in vitro* studies show that the Ser46Phe mutation in Tat results in enhanced transactivation. Next, to demonstrate how Ser46Phe contributed toward high transactivation activity. Wild-type TatC (that lacked Ser46Phe) was used as a reference Tat to compare their differential abilities to interact with TAR and their intracellular protein stability. We also used two other subtype C variants, TatN12 and TatVT6 (that lacked Ser46Phe) found in the studied population. These variants showed less transactivation than TatD60 (with Ser46Phe).

Tat transactivates HIV-1 LTR by binding to TAR, which is a critical step in the process of transcription (Huq et al., 1997; Rana and Jeang, 1999). Residues vital for Tat function include Lys28, Lys41, Lys50, Lys51, and Lys71 (for acetylation); Arg57 and Arg56 (for TAR interaction); and Tyr47, Cys22, Cys31, and Cys34 (for LTR transactivation) (Huo et al., 2011). Most of these residues were highly conserved among North Indians; however, we found that variants with Ser46Phe showed higher transactivation of LTR than TatC indicating the importance of this mutation. We studied 15 Tat variants for their ability to induce transactivation and found that variants with similar mutations resulted in similar levels of transactivation indicating the vital role of genetic variations in modulating functions. From 15 variants, we selected three variants namely TatN12 (Leu35Pro; Gly44Ser), TatD60 (Ser46Phe), and TatVT6 (B/C recombinant) as representative variants to understand the role of genetic determinants on transactivation, TAR interaction and stability. TatVT6 showed increased transactivation, which may be due to subtype C specific changes in the N-terminus and subtype B specific changes in the C-terminus. TatN12 showed less transactivation (not significant) than wild-type TatC which could possibly be due to unique mutations in TatN12 leading to weak interaction with TAR. However, the biological and clinical relevance of the reported Tat mutations remains to be established with reference to TAR sequence variations from patients.

MD simulation is a useful technique to determine the stability of biological complexes (Hornak et al., 2006; Deng et al., 2011; Johnson et al., 2012; Venken et al., 2012; Zhao et al., 2013; Vijayan et al., 2014), here, we utilized this technique to calculate the fluctuations and the binding free energy of Tat–TAR complex (Nifosi et al., 2000). Data generated from MD simulation substantiate our *in vitro* studies and provide further insights into a stronger interaction of TatD60 with TAR. The TatD60–TAR complex was more stable during the simulation with a stronger interaction between Tat and TAR. As predicted in the electrophoretic mobility shift assay (EMSA) experiment, the interaction between TatD60 (with Ser46Phe) and TAR were different in the simulation experiment. This difference has led to additional H-bond interactions between the residues namely Tyr26, Tyr29, Cys30, Ser31, Tyr47, and Ser70 in TatD60 to TAR interaction, whereas wild-type TatC lacks these residue interactions toward TAR. In case of TAR, ribonucleotide

U23 and A22 led to additional H-bonds interaction with TatD60, resulting in strong TAR interactions. Further, binding free energy calculation and residue-wise energy decomposition analyses clearly suggested that TatD60–TAR complex interacted more stably when compared to wild-type TatC–TAR complex. Particularly, the residues Gln54 and Asn67 in TatD60, but not in TatC, contributed to the binding.

TatD60 appeared to have a more stable half-life than other TatC variants within the host cells, possibly due to lower level of ubiquitination, it is clear from the fact that it has enhanced transactivation. Given its enhanced potency for transactivation and its higher stability in the cells, TatD60 has the potential to be more virulent than other TatC variants studied here. This hypothesis was supported by our observation of CD4 counts in the patients that were the source of these variants (Supplementary Table S3). However, these predictions need to be tested in a large sample population. Correlating Tat variants, especially variants with Ser46Phe, with the course of the disease, as well as response to ART, would provide molecular and clinical insights to managing HIV-1 patients in North India.

There could be other factors involved in the decline of CD4 counts in HIV-1 infected patients, however, we believe that this study with the control groups (that lacked Ser46Phe) showcase the genetic variations of Tat that could be one of the reasons for the decline. At present, it was difficult to draw a conclusion on CD4 counts and viral loads with the small sample size and warrant further need for monitoring the genetic evolution of HIV-1 strains among North Indians.

Future studies to evaluate the binding efficiency of Tat variants during complex with P-TEFb (Schulze-Gahmen et al., 2014) should further support this relationship. Taken together, this study illustrates the importance of point mutations for modulating the specific functional activities of Tat, which include increasing transactivation levels. These variants have the ability to emerge as a virulent HIV-1 strains in North India. Retrospective studies on Tat–TAR complex show that inhibition of this complex is an attractive target for developing novel antiviral drugs (Yang, 2005; Mousseau et al., 2015) and Tat could be also used as a vaccine candidate (Ensoli et al., 2016); therefore, manipulation of Tat–TAR interaction by silencing important residues is critical in forming the Tat–TAR complex as this would provide a strategy for suppressing viral gene expression. Continued studies are needed to elucidate how Tat variants manipulate the host immune cells in this population. Further, our results suggest that the nucleotides of Tat and their functions should be routinely investigated as a correlation of transactivation levels. Thus, this study provides valuable insights into the evolving events underlying the ability of the virus to adapt and enhance its replication by generating mutations and recombination events.

MATERIALS AND METHODS

Ethics Statement and Collection of Samples

This study design was approved by Research Project Advisory Committee, Institutional Biosafety Committee, and Institutional Ethical Committee for Human Research of University College of Medical Sciences (UCMS) and Guru Teg Bahadur (GTB) Hospital, Delhi, India, and from Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India. These institutes are mentored by the National AIDS Control Organization (NACO), Ministry of Health and Family Welfare, Government of India that provides free ART to HIV-1 seropositive patients under a structured HIV/AIDS Control Program. Written informed consent was obtained from HIV-1 infected adult patients ($n = 105$) and from the guardians of HIV-1 infected children participants ($n = 15$) in this study. Blood samples were collected from HIV-1 infected patients ($n = 120$; males = 68, females = 52) registered and monitored at immunodeficiency clinics in GTB Hospital and PGIMER during the period from 2004 to 2010.

Estimation of CD4 Counts and Viral Loads

HIV-1 infects CD4+ T cells and reduces these cells to a minimum level which is an indicator for the early risk of acquired immunodeficiency syndrome (AIDS) than viral load. People living with HIV AIDS (PLHA) patients under ART have shown improved CD4 counts and longer life span. ART was started in HIV-1 infected patients having CD4 count below 350/ml and children with varying counts. In this view, our study was undertaken to evaluate the correlation between Tat genetic variations on CD4 counts after 6 months on ART. From ART patients, blood was collected (CD4 counts were measured at this stage) and Tat gene was amplified, after 6 months on ART, once again CD4 counts and viral loads were measured and analyzed statistically. The CD4 count was estimated by flow cytometry using manufacturer's directions (BD Biosciences) and viral load was measured using Real Time PCR (Taqman). The viral load of less than 50 copies/ml was defined as a viral suppression or undetectable viral load.

DNA Isolation and Polymerase Chain Reaction (PCR)

Genomic DNA was extracted from PBMCs of HIV-1 infected patients by QIAamp DNA Blood Mini Kit (Qiagen) and HIV-1 subtype C (Indian isolate 93IN905 GenBank accession number AF067158 obtained from AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH) were used for amplification by PCR using the following primers:

Forward primer: 5′-ATGGAGCCAGTAGATCCTAACCTA-3′
Reverse primer: 5′-TTGCTTTGATATAAGATTTTGATGA TCCT-3′

PCR was carried out in a 15 μl reaction volume. The reaction mixture contained 500 ng genomic DNA (2 μl), 10× PCR Buffer (1.5 μl), 10 mM dNTP mix (0.37 μl), 1 μL of each primer (25 pmol), 0.25 μl of Takara Taq DNA polymerase and 8.88 μl of DNase/RNase free water. PCR conditions for the above primer sets were as follow: initial denaturation at 94°C for 5 min (1 cycle), 30 cycles of denaturation at 94°C for 15 s, annealing at 63°C for 30 s and extension at 72°C for 40 s, and

a final extension at 72°C for 5 min (1 cycle). PCR amplified products were analyzed on 1.5% agarose gel. Tat amplified from Indian HIV-1 subtype C (C.IN.93.93IN905) was used as wild-type TatC for comparison study with Tat variants. The Tat sequences were amplified both from the PBMCs viral DNA and also from the plasma viral RNA to find the exact genetic variations. The nucleotide sequences of Tat from both the source were similar in the pattern.

Cloning, Sequencing, and HIV-1 Sub-typing

The gel purified PCR products were cloned into pGEM-T easy vector (Promega). The ligation reaction was incubated at 4°C for 10 h then the ligation mix was added to LB ampicillin plates with *E. coli* DH5α strain. The plates were incubated overnight at 37°C. The positive clones were selected by picking a single colony and grown in 5 ml LB Broth with ampicillin (100 μg/ml) and incubated overnight at 37°C. Plasmid DNA was isolated from the culture by QIAprep Spin Mini Kit (Qiagen). The positive clones were screened by restriction digestion of plasmid DNA with *EcoRI* in a 10 μl reaction volume at 37°C for 2 h. The digested products were analyzed on a 1.5% agarose gel. The positive clones were commercially sequenced from LabIndia and SciGenom laboratories. The nucleotide sequences were assembled and error was checked by using BLAST to search for sequence similarities to previously reported sequences in the databases and to eliminate potential laboratory errors. HIV-1 sub-typing, recombination, phylogenetic tree and mutational analyses have been carried out as described (Ronsard et al., 2014, 2015).

Plasmids and Antibodies

TatC (lack Ser46Phe) and Tat variants (TatN12, TatD60, TatVT6) were cloned into: (a) mammalian expression vector pCMV-myc vector (Clonetech) under the CMV promoter for functional studies, and (b) prokaryotic expression vector pGEX-4T-2 (Invitrogen) to obtain GST-tagged proteins. HIV-1 subtype B TAR was cloned in pcDNA3.1 (Invitrogen) for TAR synthesis to determine Tat–TAR binding activity. Anti-Tat antibody (NIH AIDS Reagent Programme), Anti-myc antibody (Clontech), Anti-GAPDH antibody (Cell Signaling Technology), Anti-rabbit IgG conjugated to HRP (Jackson Immunoresearch), and Anti-Mouse IgG conjugated to HRP (Jackson Immunoresearch) were used in western blotting.

Cell Culture and Transfection

Human embryonic kidney (HEK) 293 cells (NIH AIDS Reagent Programme) were maintained in Dulbecco's modified Eagle's medium (DMEM) (Himedia Laboratories) supplemented with L-glutamine and sodium pyruvate, fetal calf serum (10%), penicillin (100 U/ml), streptomycin (0.1 mg/ml), and amphotericin B (0.25 μg/ml) at 37°C in the presence of 5% CO_2. Cells were transfected with lipofectamine 2000 (Invitrogen) in serum free DMEM media.

Western Blotting

HEK293 cells were transfected with 1 μg of pCMV-myc Tat variants (TatN12, TatVT6, and TatD60) and wild-type TatC. After 24 h, cells were harvested and total protein was extracted using RIPA lysis buffer (Invitrogen). The amount of protein was estimated by BCA Assay (Pierce). Tat proteins were run on 12% SDS-PAGE and transferred to the nitrocellulose membrane (BIORAD) using standard methods (Chaudhuri et al., 2015; Mohankumar et al., 2015; Sridharan et al., 2015). The membrane was incubated with anti-myc antibody followed by Anti-rabbit IgG conjugated to HRP. The membrane was developed using ECL reagent (Amersham). GAPDH was used as a loading control and the expression of proteins was normalized with the amount of GAPDH. 1 μg of the empty pCMV-myc vector was used as a control in all the experiments and the experiment was repeated three times for confirmation of the result.

Luciferase Reporter Assay

HEK293 cells were co-transfected with 200 ng of pCMV-myc Tat variants and wild-type TatC in each well of 6 well plate along with 50 ng of pGL3-Luc vector containing subtype C LTR. Cells were transfected only with subtype C. LTR construct was used as a control. After 24 h of transfection, cells were harvested and lysed with reporter lysis buffer (Promega) and luciferase activity was measured in the luminometer. 200 ng of the empty pCMV-myc vector was used as a control and the luciferase activity was normalized to the empty vector; the experiment was performed in triplicate.

Purification of Tat Proteins

Tat variants and wild-type Tat were cloned into pGEX-4T-2 vector. *E. coli* BL21 (DE3) PlysS cells were transformed with recombinant plasmids and grown at 37°C overnight. Recombinant protein expression was induced by IPTG for 3 h at 37°C. Cells were harvested, disrupted and recombinant proteins were purified using Glutathione-agarose (Pierce) using manufacturer's directions.

Electrophoretic Mobility Shift Assay (EMSA)

Subtype B TAR was cloned between HindIII and BamHI site in pCDNA3 vector (Promega). ^{32}P-labeled TAR was transcribed *in vitro* using T7 RNA polymerase. TAR was incubated with increasing amounts of purified Tat protein (0.1–2 μg) for 10 min on ice, followed by 10 min at 37°C with binding buffer (Promega). The reaction was stopped by adding 4X gel loading buffer and Tat variants with TAR complexes were analyzed on 4% Non-denaturing polyacrylamide gels and autoradiography was done. 1 μg of the empty pCMV-myc empty vector was used as a control and the expression of Tat–TAR binding was normalized with interaction with empty vector; the experiment was repeated three times.

Cycloheximide Chase Assay

HEK293 cells were transfected with 1 μg of pCMV-myc Tat variants. After 24 h, cycloheximide was added (final

concentration 100 μg/mL). Cells were harvested at different time intervals (0, 1, 2, and 3 h). Cell lysates were made with 1X RIPA lysis buffer and resolved by 12% SDS-PAGE. Anti-Tat antibody was used for detection by immunoblotting. GAPDH was used as a loading control and the expression of proteins was normalized with the amount of GAPDH. Tat proteins expressed after 24 h of transfection (before adding cycloheximide) were used as controls; the experiment was repeated three times.

In Vitro Ubiquitination Assay

HEK293 cells were co-transfected with 1 μg of pCMV-myc Tat variants and 1 μg of His6-Ubiquitin Protein for 24 h and processed as previously described (Verma et al., 2011). The expression of proteins was normalized with the amount of empty vector. 1 μg of the empty pCMV-myc vector was used as a control; the experiment was repeated three times.

Homology Models and MD Simulations

Homology models of Tat protein variants were generated using the solution structure of Tat protein as a template (PDB ID: 1TAC) and a crystal structure (PDB ID: 5L1Z) using Modeller 9v8 (Eswar et al., 2007) and then docked using HADDOCK web server (Guru Interface) (Eisenberg et al., 1997). Models were validated using PROCHECK (Laskowski et al., 1996) and the 3D-1D score of Verify3D (Bowie et al., 1991; Luthy et al., 1992). MD simulations were performed using GROMACS v4.5.6 (Van Der Spoel et al., 2005; Pronk et al., 2013) with AMBER99SB-ILDN force field (Lindorff-Larsen et al., 2010). Tat–TAR complex was solvated in a cubic box using TIP3P water model. The solvated systems were subjected to energy minimization using steepest descent and conjugate gradient algorithms keeping energy gradient convergence cut off of 10 kJ mol^{-1} nm^{-1}. LINCS algorithm was used to calculate all the covalent bonds with hydrogen. The time step was kept at 2 femtoseconds (fs) for the simulation. The cut-off distance of 10 Å was used for all short-range non-bonded interactions and 12 Å Fourier grid spacing in PME was used for long-range electrostatics. NVT and NPT steps were run for 250 picoseconds (ps) and the final production run was done for 20 nanoseconds (ns).

Binding Free Energy Calculations

The binding free energy of Tat–TAR complex was estimated by using MM/GBSA python scripts implemented in Amber11 package (Kollman et al., 2000; Campanera and Pouplana, 2010). Energy calculations were done over 5000 frames of 5 ns trajectory. Residue-wise energy decomposition studies were performed over the same trajectory using Amber decomposition script, which highlights important interactions between Tat proteins and TAR, and to identify the crucial residues in Tat proteins.

Statistical Analysis

Data were analyzed using the SPSS 7.5-Windows student version software (SPSS, Inc., Chicago, IL, USA). One-way ANOVA followed by Tukey's test was used to assess statistical significance between groups ($P < 0.05$ represents significance and $P < 0.01$ represents high significance) (Mohankumar et al., 2014a,b).

Accession Numbers

Sequences of 120 Tat variants are available at – (GenBank: FJ432068-FJ432079, FJ210870-FJ210875, EU583126-EU583128, EU551665, FJ429357, FJ429358, HQ110624-HQ110630, HQ110608-HQ110623, JQ918787-JQ918788, GU451679-GU451681, and HQ011384-HQ011385). Sequences of unique Tat variants are available at (GenBank: TatN12 – HQ110625, TatVT6 – FJ432073, TatD60 – HQ110614).

AUTHOR CONTRIBUTIONS

LR and AB conceived and designed the experiments. LR and NG performed the experiments. VS performed simulation experiment. LR, VS, KM, TR, SS, DR, MC, and AB analyzed and interpreted the data. LR, NG, VS, KM, TR, SS, SP, BK, DR, SC, MC, VR, and AB contributed reagents/materials/analysis tools. LR, VS, TR, MC, VR, and AB wrote the manuscript. LR, KM, TR, BK, VR, and AB edited the manuscript.

FUNDING

This study was supported by Department of Biotechnology (BT/PR10599/Med/29/76/2008) and Indian Council of Medical Research (HIV/50/142/9/2011-ECD-II), Government of India, to Dr. AB, National Institute of Immunology, New Delhi, India and Dr. VR, UCMS and GTB Hospital, Delhi, India. Dr. MC gratefully acknowledges Department of Biotechnology (DBT), Government of India (DBT's Twining programme for North East-BT/246/NE/TBP/2011/77) for the purchase of server.553.

ACKNOWLEDGMENTS

We would like to thank Dr. Vidhya Vijayakumar, BCH, Boston, USA for editing the manuscript. We appreciate the help from Dr. Ajay Wanchu, PGIMER, Chandigarh, India for providing HIV-1 infected blood samples. We thank Dr. C. Ganeshkumar, IIPM, Bangalore, India for helping in the statistical analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.00706/full#supplementary-material

REFERENCES

Arya, S. K., Guo, C., Josephs, S. F., and Wong-Staal, F. (1985). Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science* 229, 69–73. doi: 10.1126/science.2990040

Blackard, J. T., Cohen, D. E., and Mayer, K. H. (2002). Human immunodeficiency virus superinfection and recombination: current state of knowledge and potential clinical consequences. *Clin. Infect. Dis.* 34, 1108–1114. doi: 10.1086/339547

Bowie, J. U., Luthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170. doi: 10.1126/science.1853201

Bucci, M. (2015). Viral mechanisms: tat modulates DAT. *Nat. Chem. Biol.* 11, 240. doi: 10.1038/nchembio.1779

Buonaguro, L., Buonaguro, F. M., Giraldo, G., and Ensoli, B. (1994). The human immunodeficiency virus type 1 Tat protein transactivates tumor necrosis factor beta gene expression through a TAR-like structure. *J. Virol.* 68, 2677–2682.

Buonaguro, L., Tornesello, M. L., and Buonaguro, F. M. (2007). Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J. Virol.* 81, 10209–10219. doi: 10.1128/JVI.00872-07

Burton, D. R., Desrosiers, R. C., Doms, R. W., Koff, W. C., Kwong, P. D., Moore, J. P., et al. (2004). HIV vaccine design and the neutralizing antibody problem. *Nat. Immunol.* 5, 233–236. doi: 10.1038/ni0304-233

Campanera, J. M., and Pouplana, R. (2010). MMPBSA decomposition of the binding energy throughout a molecular dynamics simulation of amyloid-beta (Abeta(10-35)) aggregation. *Molecules* 15, 2730–2748. doi: 10.3390/molecules15042730

Chaudhuri, S., Bhattacharya, D., Singh, M. K., Moitra, S., Ronsard, L., Ghosh, T. K., et al. (2015). Disease relevance of T11TS-induced T-cell signal transduction through the CD2-mediated calcineurin-NFAT pathway: perspectives in glioma immunotherapy. *Mol. Immunol.* 67, 256–264. doi: 10.1016/j.molimm.2015.06.002

Coffin, J., and Swanstrom, R. (2013). HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harb. Perspect. Med.* 3:a012526. doi: 10.1101/cshperspect.a012526

Deng, N. J., Zheng, W., Gallicchio, E., and Levy, R. M. (2011). Insights into the dynamics of HIV-1 protease: a kinetic network model constructed from atomistic simulations. *J. Am. Chem. Soc.* 133, 9387–9394. doi: 10.1021/ja2008032

Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., et al. (1989). Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6925–6929. doi: 10.1073/pnas.86.18.6925

Dougherty, J. P., and Temin, H. M. (1988). Determination of the rate of base-pair substitution and insertion mutations in retrovirus replication. *J. Virol.* 62, 2817–2822.

Eisenberg, D., Luthy, R., and Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 277, 396–404. doi: 10.1016/S0076-6879(97)77022-8

Ensoli, B., Cafaro, A., Monini, P., Marcotullio, S., and Ensoli, F. (2014). Challenges in HIV vaccine research for treatment and prevention. *Front. Immunol.* 5:417. doi: 10.3389/fimmu.2014.00417

Ensoli, B., Nchabeleng, M., Ensoli, F., Tripiciano, A., Bellino, S., Picconi, O., et al. (2016). HIV-Tat immunization induces cross-clade neutralizing antibodies and CD4(+) T cell increases in antiretroviral-treated South African volunteers: a randomized phase II clinical trial. *Retrovirology* 13, 34. doi: 10.1186/s12977-016-0261-1

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., et al. (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* 54, 5.6.1–5.6.37. doi: 10.1002/0471140864.ps0209s50

Fulcher, A. J., and Jans, D. A. (2003). The HIV-1 Tat transactivator protein: a therapeutic target? *IUBMB Life* 55, 669–680.

Geretti, A. M. (2006). HIV-1 subtypes: epidemiology and significance for HIV management. *Curr. Opin. Infect. Dis.* 19, 1–7. doi: 10.1097/01.qco.0000200293.45532.68

Gupta, N., and Banerjea, A. C. (2009). C-terminal half of HIV-1 Vif C possesses major determinant for APOBEC3G degradation. *AIDS* 23, 141–143. doi: 10.1097/QAD.0b013e32831c8b08

Hamy, F., Felder, E. R., Heizmann, G., Lazdins, J., Aboul-ela, F., Varani, G., et al. (1997). An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3548–3553. doi: 10.1073/pnas.94.8.3548

Hamy, F., Gelus, N., Zeller, M., Lazdins, J. L., Bailly, C., and Klimkait, T. (2000). Blocking HIV replication by targeting Tat protein. *Chem. Biol.* 7, 669–676. doi: 10.1016/S1074-5521(00)00012-0

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., and Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126. doi: 10.1038/373123a0

Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006). HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* 103, 915–920. doi: 10.1073/pnas.0508452103

Huo, L., Li, D., Sun, X., Shi, X., Karna, P., Yang, W., et al. (2011). Regulation of Tat acetylation and transactivation activity by the microtubule-associated deacetylase HDAC6. *J. Biol. Chem.* 286, 9280–9286. doi: 10.1074/jbc.M110.208884

Huq, I., Wang, X., and Rana, T. M. (1997). Specific recognition of HIV-1 TAR RNA by a D-Tat peptide. *Nat. Struct. Biol.* 4, 881–882. doi: 10.1038/nsb1197-881

Jetzt, A. E., Yu, H., Klarmann, G. J., Ron, Y., Preston, B. D., and Dougherty, J. P. (2000). High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* 74, 1234–1240. doi: 10.1128/JVI.74.3.1234-1240.2000

Johnson, B. C., Metifiot, M., Pommier, Y., and Hughes, S. H. (2012). Molecular dynamics approaches estimate the binding energy of HIV-1 integrase inhibitors and correlate with in vitro activity. *Antimicrob. Agents Chemother.* 56, 411–419. doi: 10.1128/AAC.05292-11

Kirchhoff, F. (2009). Is the high virulence of HIV-1 an unfortunate coincidence of primate lentiviral evolution? *Nat. Rev. Microbiol.* 7, 467–476. doi: 10.1038/nrmicro2111

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33, 889–897. doi: 10.1021/ar000033j

Konings, F. A., Burda, S. T., Urbanski, M. M., Zhong, P., Nadas, A., and Nyambi, P. N. (2006). Human immunodeficiency virus type 1 (HIV-1) circulating recombinant form 02_AG (CRF02_AG) has a higher in vitro replicative capacity than its parental subtypes A and G. *J. Med. Virol.* 78, 523–534. doi: 10.1002/jmv.20572

Lalonde, M. S., Lobritz, M. A., Ratcliff, A., Chamanian, M., Athanassiou, Z., Tyagi, M., et al. (2011). Inhibition of both HIV-1 reverse transcription and gene expression by a cyclic peptide that binds the Tat-transactivating response element (TAR) RNA. *PLoS Pathog.* 7:e1002038. doi: 10.1371/journal.ppat.1002038

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486. doi: 10.1007/BF00228148

Li, L., Dahiya, S., Kortagere, S., Aiamkitsumrit, B., Cunningham, D., Pirrone, V., et al. (2012). Impact of Tat genetic variation on HIV-1 disease. *Adv. Virol.* 2012:123605. doi: 10.1155/2012/123605

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78, 1950–1958. doi: 10.1002/prot.22711

Luthy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* 356, 83–85. doi: 10.1038/356083a0

Mediouni, S., Marcondes, M. C., Miller, C., McLaughlin, J. P., and Valente, S. T. (2015). The cross-talk of HIV-1 Tat and methamphetamine in HIV-associated neurocognitive disorders. *Front. Microbiol.* 6:1164. doi: 10.3389/fmicb.2015.01164

Mohankumar, K., Pajaniradje, S., Sridharan, S., Singh, V. K., Ronsard, L., Banerjea, A. C., et al. (2014a). Mechanism of apoptotic induction in human breast cancer cell, MCF-7, by an analog of curcumin in comparison with curcumin–an in vitro and in silico approach. *Chem. Biol. Interact.* 210, 51–63. doi: 10.1016/j.cbi.2013.12.006

Mohankumar, K., Pajaniradje, S., Sridharan, S., Singh, V. K., Ronsard, L., Banerjea, A. C., et al. (2014b). Apoptosis induction by an analog of curcumin (BDMC-A) in human laryngeal carcinoma cells through intrinsic and extrinsic pathways. *Cell. Oncol. (Dordr)* 37, 439–454. doi: 10.1007/s13402-014-0207-3

Mohankumar, K., Sridharan, S., Pajaniradje, S., Singh, V. K., Ronsard, L., Banerjea, A. C., et al. (2015). BDMC-A, an analog of curcumin, inhibits markers of invasion, angiogenesis, and metastasis in breast cancer cells via NF-kappaB pathway–A comparative study with curcumin. *Biomed. Pharmacother.* 74, 178–186. doi: 10.1016/j.biopha.2015.07.024

Mousseau, G., Mediouni, S., and Valente, S. T. (2015). Targeting HIV transcription: the quest for a functional cure. *Curr. Top. Microbiol. Immunol.* 389, 121–145. doi: 10.1007/82_2015_435

Muesing, M. A., Smith, D. H., and Capon, D. J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. *Cell* 48, 691–701. doi: 10.1016/0092-8674(87)90247-9

Neogi, U., Sood, V., Ronsard, L., Singh, J., Lata, S., Ramachandran, V. G., et al. (2011). Genetic architecture of HIV-1 genes circulating in north India & their functional implications. *Indian J. Med. Res.* 134, 769–778. doi: 10.4103/0971-5916.92624

Nifosi, R., Reyes, C. M., and Kollman, P. A. (2000). Molecular dynamics studies of the HIV-1 TAR and its complex with argininamide. *Nucleic Acids Res.* 28, 4944–4955. doi: 10.1093/nar/28.24.4944

Nunnari, G., Smith, J. A., and Daniel, R. (2008). HIV-1 Tat and AIDS-associated cancer: targeting the cellular anti-cancer barrier? *J. Exp. Clin. Cancer Res.* 27:3. doi: 10.1186/1756-9966-27-3

Okamoto, H., Sheline, C. T., Corden, J. L., Jones, K. A., and Peterlin, B. M. (1996). Trans-activation by human immunodeficiency virus Tat protein requires the C-terminal domain of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11575–11579. doi: 10.1073/pnas.93.21.11575

Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., et al. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29, 845–854. doi: 10.1093/bioinformatics/btt055

Rana, T. M., and Jeang, K. T. (1999). Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.* 365, 175–185. doi: 10.1006/abbi.1999.1206

Rodriguez, M. A., Ding, M., Ratner, D., Chen, Y., Tripathy, S. P., Kulkarni, S. S., et al. (2009). High replication fitness and transmission efficiency of HIV-1 subtype C from India: implications for subtype C predominance. *Virology* 385, 416–424. doi: 10.1016/j.virol.2008.12.025

Romani, B., Engelbrecht, S., and Glashoff, R. H. (2010). Functions of Tat: the versatile protein of human immunodeficiency virus type 1. *J. Gen. Virol.* 91, 1–12. doi: 10.1099/vir.0.016303-0

Ronsard, L., Lata, S., Singh, J., Ramachandran, V. G., Das, S., and Banerjea, A. C. (2014). Molecular and genetic characterization of natural HIV-1 Tat Exon-1 variants from North India and their functional implications. *PLoS ONE* 9:e85452. doi: 10.1371/journal.pone.0085452

Ronsard, L., Raja, R., Panwar, V., Saini, S., Mohankumar, K., Sridharan, S., et al. (2015). Genetic and functional characterization of HIV-1 Vif on APOBEC3G degradation: first report of emergence of B/C recombinants from North India. *Sci. Rep.* 5:15438. doi: 10.1038/srep15438

Santoro, M. M., and Perno, C. F. (2013). HIV-1 genetic variability and clinical implications. *ISRN Microbiol.* 2013:481314. doi: 10.1155/2013/481314

Schulze-Gahmen, U., Lu, H., Zhou, Q., and Alber, T. (2014). AFF4 binding to Tat-P-TEFb indirectly stimulates TAR recognition of super elongation complexes at the HIV promoter. *Elife* 3:e02375. doi: 10.7554/eLife.02375

Sharp, P. M., and Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harb. Perspect. Med.* 1:a006841. doi: 10.1101/cshperspect.a006841

Siu, K. K., Sultana, A., Azimi, F. C., and Lee, J. E. (2013). Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nat. Commun.* 4, 2593. doi: 10.1038/ncomms3593

Sood, V., Ranjan, R., and Banerjea, A. C. (2008). Functional analysis of HIV-1 subtypes B and C HIV-1 Tat exons and RGD/QGD motifs with respect to Tat-mediated transactivation and apoptosis. *AIDS* 22, 1683–1685. doi: 10.1097/QAD.0b013e3282f56114

Sridharan, S., Mohankumar, K., Jeepipalli, S. P., Sankaramourthy, D., Ronsard, L., Subramanian, K., et al. (2015). Neuroprotective effect of *Valeriana wallichii* rhizome extract against the neurotoxin MPTP in C57BL/6 mice. *Neurotoxicology* 51, 172–183. doi: 10.1016/j.neuro.2015.10.012

Sun, G., Li, H., Wu, X., Covarrubias, M., Scherer, L., Meinking, K., et al. (2012). Interplay between HIV-1 infection and host microRNAs. *Nucleic Acids Res.* 40, 2181–2196. doi: 10.1093/nar/gkr961

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701–1718. doi: 10.1002/jcc.20291

Venken, T., Daelemans, D., De Maeyer, M., and Voet, A. (2012). Computational investigation of the HIV-1 Rev multimerization using molecular dynamics simulations and binding free energy calculations. *Proteins* 80, 1633–1646. doi: 10.1002/prot.24057

Verma, S., Ali, A., Arora, S., and Banerjea, A. C. (2011). Inhibition of beta-TrcP-dependent ubiquitination of p53 by HIV-1 Vpu promotes p53-mediated apoptosis in human T cells. *Blood* 117, 6600–6607. doi: 10.1182/blood-2011-01-333427

Vijayan, R. S., Arnold, E., and Das, K. (2014). Molecular dynamics study of HIV-1 RT-DNA-nevirapine complexes explains NNRTI inhibition and resistance by connection mutations. *Proteins* 82, 815–829. doi: 10.1002/prot.24460

Wang, X., Herr, R. A., Chua, W. J., Lybarger, L., Wiertz, E. J., and Hansen, T. H. (2007). Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. *J. Cell Biol.* 177, 613–624. doi: 10.1083/jcb.200611063

Wolinsky, S. M., Korber, B. T., Neumann, A. U., Daniels, M., Kunstman, K. J., Whetsell, A. J., et al. (1996). Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* 272, 537–542. doi: 10.1126/science.272.5261.537

Yang, M. (2005). Discoveries of Tat-TAR interaction inhibitors for HIV-1. *Curr. Drug Targets Infect. Disord.* 5, 433–444. doi: 10.2174/156800505774912901

Yuan, Y., Huang, X., Midde, N. M., Quizon, P. M., Sun, W. L., Zhu, J., et al. (2015). Molecular mechanism of HIV-1 Tat interacting with human dopamine transporter. *ACS Chem. Neurosci.* 6, 658–665. doi: 10.1021/acschemneuro.5b00001

Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., et al. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* 497, 643–646. doi: 10.1038/nature12162

Zhou, C., and Rana, T. M. (2002). A bimolecular mechanism of HIV-1 Tat protein interaction with RNA polymerase II transcription elongation complexes. *J. Mol. Biol.* 320, 925–942. doi: 10.1016/S0022-2836(02)00556-9

Zhu, C., Gao, W., Zhao, K., Qin, X., Zhang, Y., Peng, X., et al. (2013). Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. *Nat. Commun.* 4, 2722. doi: 10.1038/ncomms3722

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., et al. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev.* 11, 2622–2632. doi: 10.1101/gad.11.20.2622

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ronsard, Ganguli, Singh, Mohankumar, Rai, Sridharan, Pajaniradje, Kumar, Rai, Chaudhuri, Coumar, Ramachandran and Banerjea. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

REVIEW
published: 09 January 2017
doi: 10.3389/fmicb.2016.02152

Unusual Fusion Proteins of HIV-1

*Simon Langer and Daniel Sauter**

Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany

Despite its small genome size, the Human Immunodeficiency Virus 1 (HIV-1) is one of the most successful pathogens and has infected more than 70 million people worldwide within the last decades. In total, HIV-1 expresses 16 canonical proteins from only nine genes within its 10 kb genome. Expression of the structural genes *gag*, *pol*, and *env*, the regulatory genes *rev* and *tat* and the accessory genes *vpu*, *nef*, *vpr*, and *vif* enables assembly of the viral particle, regulates viral gene transcription, and equips the virus to evade or counteract host immune responses. In addition to the canonically expressed proteins, a growing number of publications describe the existence of non-canonical fusion proteins in HIV-1 infected cells. Most of them are encoded by the *tat-env-rev* locus. While the majority of these fusion proteins (e.g., TNV/p28^{tev}, p18^{6Drev}, Tat1-Rev2, Tat^8c, p17^{tev}, or Ref) are the result of alternative splicing events, Tat-T/Vpt is produced upon programmed ribosomal frameshifting, and a Rev1-Vpu fusion protein is expressed due to a nucleotide polymorphism that is unique to certain HIV-1 clade A and C strains. A better understanding of the expression and activity of these non-canonical viral proteins will help to dissect their potential role in viral replication and reveal how HIV-1 optimized the coding potential of its genes. The goal of this review is to provide an overview of previously described HIV-1 fusion proteins and to summarize our current knowledge of their expression patterns and putative functions.

Keywords: HIV-1, fusion protein, gene fusion, alternative splicing, polymorphism, ribosomal frameshift

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Kei Sato,
Kyoto University, Japan
Michael M. Thomson,
Instituto de Salud Carlos III, Spain

****Correspondence:***
Daniel Sauter
daniel.sauter@uni-ulm.de

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *28 October 2016*
Accepted: *20 December 2016*
Published: *09 January 2017*

Citation:
Langer S and Sauter D (2017) Unusual Fusion Proteins of HIV-1. Front. Microbiol. 7:2152. doi: 10.3389/fmicb.2016.02152

INTRODUCTION

The genome of the Human Immunodeficiency Virus type 1 (HIV-1), the major causative agent of the current AIDS pandemic, is relatively small, comprising <10,000 bases in total. Arranged in three different reading frames, it contains only nine canonical genes (**Figure 1**). Nevertheless, the virus replicates and spreads efficiently in its human host, which expresses about 2500 times more protein-coding genes from a three billion base pair genome. How does a retrovirus with its limited genome size manage to keep pace in this David vs. Goliath struggle? How can such a tiny genome encode all the tools that are required for efficient replication and immune evasion in such a hostile environment? One major advantage of HIV-1 and related retroviruses compared to their host species is certainly their high mutation rate that allows them to quickly adapt to an ever-changing environment. Furthermore, viral proteins are often multifunctional and exert a multitude of immune evasion activities. The paragon of such a multitasking or moonlighting protein is HIV-1 Nef, which has been described to downmodulate a variety of surface receptors including CD4, MHC class I, CD28, and CXCR4, counteracts the host restriction factors SERINC3/5, and upregulates the invariant chain/CD74 to suppress antigen presentation (Pereira and daSilva, 2016). Finally, viral genomes are often very compact, containing overlapping genes that encode for bi- and multi-cistronic mRNAs. As a result, viruses frequently utilize non-canonical translation mechanisms such

as internal ribosomal entry, leaky scanning, ribosomal frameshifting, shunting, or reinitiation (Firth and Brierley, 2012). Another important mechanism increasing the coding capacity of viral genomes is alternative splicing. HIV-1 and related lentiviruses contain dozens of splice donor and acceptor sites that allow the generation of more than 100 different mRNA species (Ocwieja et al., 2012; **Figure 1**). Generation and translation of these mRNAs are tightly regulated throughout the viral replication cycle and enable the coordinated synthesis of structural, regulatory and accessory proteins in an optimized ratio. For example, expression of the HIV-1 regulatory proteins Tat and Rev requires the joining of two exons (*tat1/2* or *rev1/2*) via splicing at donor D4 and acceptor A7, whereas all four accessory proteins (Vif, Vpr, Vpu, and Nef) are encoded by mono-exonic genes. Notably, *vpu* overlaps with the viral envelope (*env*) gene and both are expressed from bicistronic mRNA species (**Figure 1**). Translation of downstream *env* is enabled by a weak Kozak sequence of *vpu* (leaky scanning) and/or ribosomal shunting mechanisms that allow to bypass upstream AUG codons (Anderson et al., 2007). While expression of Env as well as all accessory and regulatory proteins requires splicing, Gag and Pol are encoded by the 5′ half of the viral

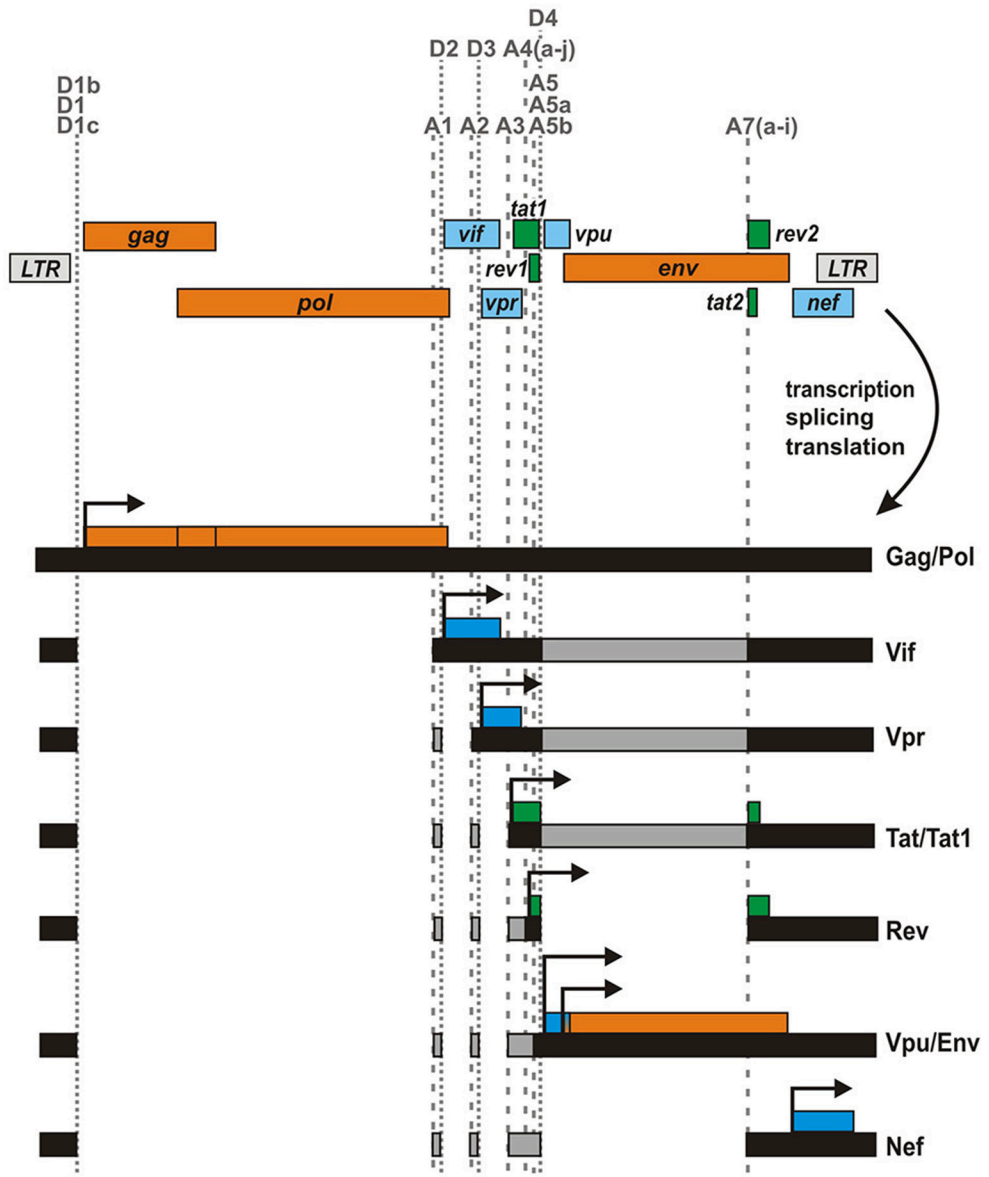


FIGURE 1 | Genome and major mRNA transcripts of HIV-1. The HIV-1 genome comprises nine canonical genes that are arranged in three different reading frames. These genes encode for structural/enzymatic (orange), regulatory (green), and accessory (blue) proteins and are flanked by two long terminal repeats (LTR). The viral mRNA is spliced into more than 100 different mono- or multi-cistronic mRNA transcripts (Ocwieja et al., 2012), encoding different viral proteins. Splice donor (D) and acceptor (A) sites are indicted by dotted and dashed vertical lines, respectively.

genome and expressed by unspliced viral mRNA. Gag can either be expressed alone or, upon ribosomal frameshifting, as a Gag-Pol poly-protein. The three precursor proteins Gag, Gag-Pol, and Env are proteolytically processed into mature proteins: Gag is cleaved by the viral protease into matrix, capsid, nucleocapsid, and the p6 protein. Similarly, the viral protease generates the mature viral enzymes reverse transcriptase (p51 and p66), protease and integrase from the Pol precursor protein. Finally, the Envelope protein is cleaved by the cellular protease furin into its mature subunits gp120 and gp41.

Considering the multitude of modulatory processes underlying the expression of viral proteins, it is not surprising that several studies have reported the expression of non-canonical fusion proteins by HIV-1. While the majority of these fusion proteins are the result of alternative splicing events joining regular or cryptic open reading frames, two of them are expressed only upon ribosomal frameshifting or gene rearrangements, respectively. The aim of this review article is to provide an overview of previously described fusion proteins of HIV-1. We will summarize our current knowledge on their expression and generation by different HIV-1 strains, discuss possible roles during the retroviral life cycle and critically review a potential relevance for viral replication.

HIV-1 FUSION PROTEINS GENERATED BY ALTERNATIVE SPLICING

TNV/p28^{tev}

In addition to the canonical splice sites, several studies have reported the presence of alternative or cryptic splice sites in different clades of HIV-1 group M (Purcell and Martin, 1993; Ocwieja et al., 2012; Vega et al., 2016). While some of these sites are conserved among diverse HIV-1 isolates and seem to be used regularly, others have only been identified in single clones of HIV-1 and/or become only active upon mutation of canonical splice sites. Two well-described examples for cryptic sites are splice acceptor 6 (A6) and donor 5 (D5), which have been identified in the genomes of HXB2 and closely related subtype B strains (**Figure 2**). Utilization of these sites results in the generation of a small exon (116 bases) derived from the *env* open reading frame (ORF). cDNA analyses revealed that this exon (designated 6D) may be fused to *tat1* and *rev2* encoding exons via splice donor 4 (D4) and acceptor 7 (A7), respectively (Feinberg et al., 1986; Wright et al., 1986; Benko et al., 1990; Salfeld et al., 1990; Schwartz et al., 1990a).

Two groups demonstrated that the respective mRNA can be translated into an unusual tripartite fusion protein comprising Tat1, 38 amino acids of Env including its V1 loop, and Rev2 (Benko et al., 1990; Salfeld et al., 1990). Salfeld and colleagues found that this protein migrates at an apparent size of 26 kDa and was probably identical to the Tat-related protein p26 described in earlier studies (Feinberg et al., 1986; Wright et al., 1986). In reference to its parental proteins Tat, Env, and Rev, the fusion protein was named TNV (Salfeld et al., 1990). The group of Barbara Felber identified the same protein independently and termed it p28^{tev} as it migrated slightly slower in their experiments (Benko et al., 1990). Both studies analyzed the closely related HIV-1 M clade B HXB2 and/or HXB3 clones. While Salfeld et al. analyzed the expression of the fusion protein only in transfected COS-7 cells, Benko and colleagues demonstrated that TNV/p28^{tev} is also expressed in various human T cell lines infected with HIV-1 HXB2 (**Table 1**). In agreement with the finding that the N-terminus of Tat, encoded by *tat1*, is sufficient to transactivate viral transcription (Sodroski et al., 1985; Cullen, 1990; Vives et al., 1994), TNV/p28^{tev} also enhances LTR-mediated gene expression and may thus represent a *bona fide* regulatory protein. Reporter assays revealed that the transactivating activity of TNV/p28^{tev} is only about 30% lower than that of its parental Tat protein (Benko et al., 1990; Salfeld et al., 1990). In contrast, the fusion protein exerts no (Salfeld et al., 1990) or only weak (Benko et al., 1990) Rev activity. Due to its chimeric structure, TNV/p28^{tev} possesses two stretches of basic amino acids in its Tat1 and Rev2 domains that mediate nucleolar localization (Benko et al., 1990). The nuclear localization and absence of a signal peptide probably also prevents glycosylation despite the presence of four N-linked glycosylation sites in the Env-derived fragment (Salfeld et al., 1990). Radiolabeling revealed that TNV/p28^{tev} is weakly phosphorylated, probably at two phosphate acceptors near its C-terminus (Benko et al., 1990). This finding is in agreement with the observation that the fusion protein migrates as a doublet of closely spaced bands in SDS gels (Göttlinger et al., 1992).

To investigate the importance of TNV/p28^{tev} for viral replication, Göttlinger and colleagues mutated the A6 and D5 splice sites in *env* without altering its primary amino acid sequences. Experiments in Jurkat T cells and PBMCs revealed that the A6 mutant of HIV-1 HXBc2 replicated as efficiently as the respective wild type control (Göttlinger et al., 1992). These findings demonstrate that expression of TNV/p28^{tev} has no significant effect on viral replication, at least *in vitro*. Interestingly, the D5 mutant was replication-defective. However, this phenotype could be ascribed to the utilization of another cryptic splice donor that resulted in detrimental intron removal and possibly reduced Tat and Rev expression levels (Göttlinger et al., 1992). These results are in agreement with the observation that most HIV-1 strains lack the cryptic splice sites generating exon 6D (Göttlinger et al., 1992). In fact, mutations that increase the amount of TNV/p28^{tev} encoding mRNAs may be detrimental for viral replication as the expression of functional Rev is reduced (Göttlinger et al., 1992; Wentz et al., 1997).

Tat1-Rev2 (p21, p24) and Rev1-Tat2 Chimeras

Besides the TNV/p28^{tev} fusion protein, Salfeld and colleagues observed the expression of two additional proteins (p21, p24) that are detected by both Rev- and Tat-specific antisera (Salfeld et al., 1990). They hypothesized that at least one of these two proteins may represent an alternative Tat-Rev fusion product that is expressed if *tat1* is fused in frame to *rev2,* without any additional *env* sequences. Notably, comprehensive analyses of mRNA species in HIV-1 infected cells identified several neighboring splice acceptor sites at the 5′ end of *rev2/tat2* that introduce a frameshift and may result in the expression of various

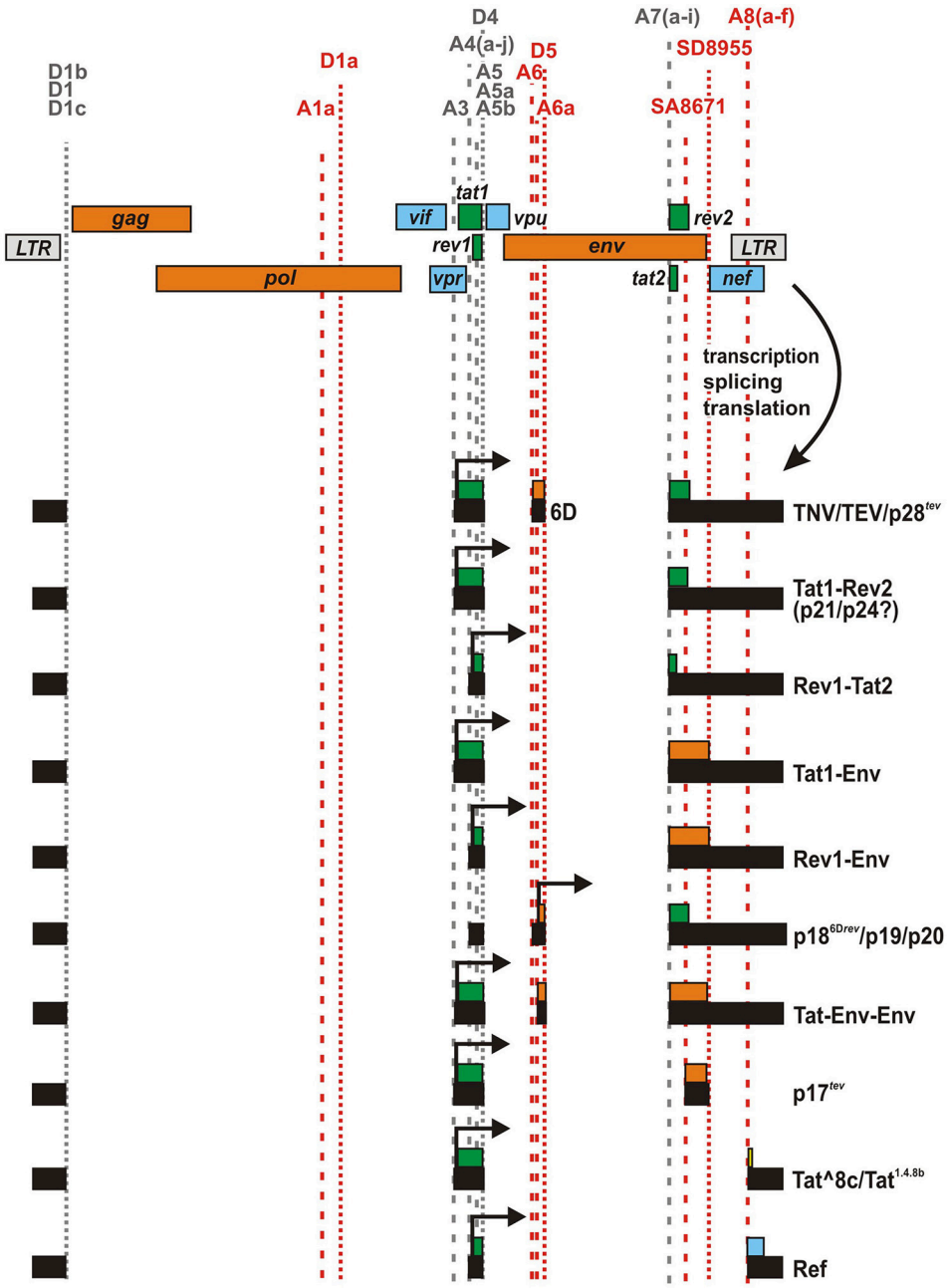


FIGURE 2 | HIV-1 fusion proteins generated by alternative splicing. Several HIV-1 strains encode cryptic (red) and/or alternative splice sites. Splicing at these sites may result in the generation of mRNA species expressing unusual fusion proteins. For example, TNV, Tat1-Rev2, Tat1-Env, Rev1-Env, p18^{6Drev}, Tat-Env-Env, p17^{tev}, and Ref comprise (parts of) Tat1, Env, Rev2, and/or Nef that are fused together in frame. In case of Tat^8c, a few amino acids encoded by the *nef*/LTR overlap (that are not in frame with the *nef* ORF) are fused to the C-terminus of Tat1.

chimeric Rev1-Tat2 or Tat1-Rev2 proteins (**Figure 2**, **Table 1**) (Schwartz et al., 1990a; Purcell and Martin, 1993; Ocwieja et al., 2012; Vega et al., 2016). Although at least some of these splice sites can be found in diverse subtypes of (primary) HIV-1 group M isolates (Vega et al., 2016), the expression of chimeric Tat/Rev proteins and their possible role in viral replication has never been investigated.

Tat1-Env and Rev1-Env Chimeras

Alternative splicing at the Rev1-Rev2/Tat1-Tat2 junction may not only result in the production of Tat-Rev chimeras, but also entail the expression of unusual Rev-Env or Tat-Env fusion proteins (**Figure 2**, **Table 1**). For example, usage of acceptors A7g and h in conjunction with donor D4 results in a +2 frameshift that enables the expression of a Tat1-Env protein (Vega et al., 2016).

TABLE 1 | HIV-1 fusion proteins.

Fusion protein	Parental proteins	Size	Function	Detected in/by			References
				Virus	Cell type	Method	
TEV/p28^{tev}	Tat1, Env, Rev2	28 kDa	Tat activity, weak Rev activity	HIV-1 M HXB2 and HXB3	Chronically infected MOLT-4/IIIB, H9/HXB2, H9/HXB3, TH4-7-5 and LC-5/HIVIIIB cells, transfected HeLa cells	IP, IF, WB, SB, sequencing	Benko et al., 1990; Schwartz et al., 1990a; Neumann et al., 1994
TNV	Tat1, Env, Rev2	26 kDa	Tat activity, no Rev activity, dispensable for replication	(Fragments of) HIV-1 M HXBc2	Transfected COS-1 and -7 cells, transfected murine CB2MX3-2 cells, reticulocyte extract	IP, cDNA phage library, IVT	Feinberg et al., 1986; Wright et al., 1986; Salfeld et al., 1990; Göttlinger et al., 1992
Tat/Rev chimeras (p21/p24?)	Tat, Rev	21, 24 kDa	Not analyzed	HIV-1 M NL4-3, HIV-1 M HXBc2 Primary HIV-1 M clade C (SPX10)	CD4+ CD25+ patient cells, infected PBMCs and HOS-CD4-CCR5 cells, transfected COS-7, and HeLa cells	IP, sequencing	Salfeld et al., 1990; Purcell and Martin, 1993; Ocwieja et al., 2012; Vega et al., 2016
Tat1-Env	Tat1, Env	∼210 aa	Not analyzed	Primary HIV-1 M clade C (SPX10)	CD4+ CD25+ patient cells	Sequencing	Vega et al., 2016
Rev1-Env	Rev1, Env	159 aa	Not analyzed	HIV-1 M 89.6	Infected primary CD4+ T cells, infected HOS-CD4-CCR5 cells	Sequencing	Ocwieja et al., 2012
p18^{6Drev}	Env, Rev2	18 kDa	No Rev activity	HIV-1 M HXB2 and pm213 L1	Chronically infected H9/HXB2, TH4-7-5 and CEM/pm213 cells, transfected HeLa cells	Sequencing, IP, IF, SB, WB,	Benko et al., 1990; Schwartz et al., 1990a; Neumann et al., 1994
p20	Env, Rev2	20 kDa	Not analyzed	HIV-1 M HXBc2	Reticulocyte extract	IVT, IP	Salfeld et al., 1990
p19	Env, Rev2	19 kDa	Not analyzed	HIV-1 M HXBc2	Transfected COS-7 cells	IP	Göttlinger et al., 1992
Tat-Env-Env	Tat1, Env	235 aa	Not analyzed	HIV-1 M 89.6	Infected primary CD4+ T cells, infected HOS-CD4-CCR5 cells	Sequencing	Ocwieja et al., 2012
p17^{tev}	Tat1, Env	17 kDa	Weak Tat activity	HIV-1 M IIIB	Transfected COS cells, Reticulocyte extract	SB, IP, IVT, sequencing	Furtado et al., 1991
Tat$^{1.4.8b}$	Tat1, Nef^{shift}	90 aa	Not analyzed	Primary HIV-1 M clade B	PBMCs from an HIV-1 infected individual	Sequencing	Carrera et al., 2010
Tat^8c	Tat1, Nef^{shift}	97 aa	Weak Tat activity	HIV-1 M 89.6	Infected primary CD4+ T cells, infected HOS-CD4-CCR5 cells	Sequencing	Ocwieja et al., 2012
Ref	Rev1, Nef	12.5 kDa	No Rev activity	HIV-1 M 89.6	Infected primary CD4+ T cells, infected HOS-CD4-CCR5 cells	WB, sequencing	Ocwieja et al., 2012
Tat-T/Vpt	Tat1, T (Tat1^{shift}, Vpu^{shift})	17 kDa	No detectable Rev or Tat activity	(Fragments of) HIV-1 M HXBc2, BH10 and BRU	Reticulocyte extract	IVT, IP	Cohen et al., 1990
Rev1-Vpu	Rev1, Vpu	∼14 kDa	Dispensable for replication	HIV-1 M clade A, C and CRF	Transfected HEK293T cells, infected SupT1 cells, and PBMCs	WB, sequencing	Kraus et al., 2010; Langer et al., 2015

IP, immunoprecipitation; IF, immunofluorescence; WB, Western blot; SB, Southern blot; IVT, in vitro translation; identical/similar proteins are separated by dashed lines.

Conversely, splicing at A7e induces a +1 frameshift and the resulting RNA species have the potential to express a Rev1-Env fusion protein (Ocwieja et al., 2012). Due to preferential usage of acceptor A5, however, the majority of mRNA species using alternative splice sites of A7 may lack the *rev1* and *tat1* initiation codons and express Nef instead (Ocwieja et al., 2012; Vega et al., 2016).

p18^{6Drev}/p19/p20

Depending on the specific splice acceptor used, the cryptic *env* exon 6D can be fused to different exons at its 5′ end. Notably, a TNV/p28^{tev} fusion protein can only be synthesized upon usage of splice acceptor 3 (A3) since alternative utilization of A4 and A5 results in a loss of the *tat1* initiation codon (**Figure 2**). In the latter case, two methionine residues in exon 6D may serve as alternative start codons and result in the expression of a 6D/Env-Rev2 fusion protein (Göttlinger et al., 1992; Neumann et al., 1994). Experiments in transfected HeLa and COS-7 cells as well as chronically infected H9 and CEM cells revealed that at least the HIV-1 M HXB2 clone and HIV-1 pm213 L1, a closely related strain, are able to express this fusion protein. According to its apparent size in the gel, this unusual viral protein has been termed p18^{6Drev} (Benko et al., 1990; Schwartz et al., 1990a; Wentz et al., 1997), p19 (Göttlinger et al., 1992), or p20 (Salfeld et al., 1990). In contrast to TNV/p28^{tev}, which is exclusively localized in nucleoli, p18^{6Drev} can be found in both nucleoli and the cytoplasm (Benko et al., 1990). Thus, the Env domain seems to affect the otherwise nuclear localization of Rev2. Furthermore, p18^{6Drev} did not display any significant Rev activity (Benko et al., 1990). This is in agreement with the finding that both the N- and C-terminal parts of Rev are required for nuclear targeting and functional activity of this regulatory protein (Malim et al., 1989).

Tat-Env-Env

Analyzing the HIV-1 clone 89.6, Ocwieja and colleagues identified another cryptic splice acceptor site (named A6a) that lies 52 bp downstream of A6 (Ocwieja et al., 2012). RNAs generated via splicing at this site are predicted to encode a tripartite Tat-Env-Env fusion protein comprising the N-terminus of Tat and two stretches (aa 145–169 and 716–853) of Env (**Figure 2**, **Table 1**). Similar to A6, however, acceptor A6a is not well conserved among different strains of HIV-1. This is also true for splice donor D5, which is required for the generation of both, TNV/p28^{tev} and Tat-Env-Env encoding RNA (Ocwieja et al., 2012).

p17^{tev}

In 1991, Furtado and colleagues identified a novel splice acceptor site (SA8671) in *env*, located 240 bases downstream of the canonical *tat*/*rev* splice acceptor A7 (**Figure 2**). As a result, the *tat1* encoding exon may be fused in frame to an exon encoding the C-terminal 58 amino acids of Env gp41 (Furtado et al., 1991). Experiments in transfected COS cells and rabbit reticulocyte extracts demonstrated that the respective mRNA indeed expressed a 17 kDa protein (named p17^{tev}) that can be detected by both Tat- and gp41-specific antibodies. RNase protection experiments, however, showed that p17^{tev} encoding mRNA is only expressed at very low levels, which may explain why the authors failed to detect this fusion protein in infected H9 cells or primary lymphocytes. Reporter assays revealed that p17^{tev} exerts only weak transactivating activity although it comprises the whole N-terminus of Tat, encoded by *tat1*. Unlike other mutated Tat proteins (Pearson et al., 1990), p17^{tev} did not exert any dominant negative effect on wild type Tat (Furtado et al., 1991).

Tat^8c/Tat$^{1.4.8b}$

In contrast to splice acceptors A6 and SA8671, which have only been detected in few lab-adapted clones of HIV-1, several groups reported the presence of additional splice acceptor sites (A8a–e) in the *nef* genes of both laboratory-adapted and primary isolates of HIV-1 (Smith et al., 1992; Carrera et al., 2010; Ocwieja et al., 2012) (**Figure 2**). These sites result in the generation of a previously unappreciated class of 1 kb transcripts. Intriguingly, A8c may be used as frequently as acceptor A7, which is required for expression of regular Rev and Tat proteins (Ocwieja et al., 2012). Splice events joining donor D5 to acceptors A8a–e result in mRNA species that have the potential to encode Tat1-Nef fusion proteins. For example, Carrera and colleagues predicted the expression of a Tat$^{1.4.8b}$ protein upon splicing of D4 to A8b. This fusion protein consists of the N-terminus of Tat and 18 amino acids encoded by the *nef*/LTR region (Carrera et al., 2010). Notably, however, the 18 C-terminal amino acids do not contain any functional motifs of Nef, as *tat$^{1.4.8b}$* and *nef* are not translated in the same reading frame. More recently, Ocwieja and colleagues identified mRNA species in infected primary CD4+ T cells, which resulted from splicing of D4 to A8c (Ocwieja et al., 2012). These mRNAs express a Tat^8c fusion protein consisting of Tat1 and 25 novel amino acids encoded by the *nef*/LTR locus. In transfected TZM-bl cells, this protein exerted only weak transactivating activity. Notably, analyses of PBMCs from HIV-1 infected individuals demonstrated that acceptors A8 may also be fused to donor D1, resulting in RNA species that have the potential to encode a truncated protein, consisting of the C-terminal 34 amino acids of Nef (Smith et al., 1992; Carrera et al., 2010). Although the initiation codon of this protein, named C-Nef-34, is conserved among most clades of HIV-1 group M (Carrera et al., 2010) and although the majority of HIV-1 and HIV-2 strains contain at least one A8 site (Ocwieja et al., 2012), the importance of C-Nef-34 and/or Tat^8c for viral replication has remained unclear.

Ref

Tat^8c/Tat$^{1.4.8b}$ is not the only fusion protein containing amino acid sequences encoded by the *nef*/LTR locus of HIV-1. cDNA sequence analyses of cells infected with HIV-1 M 89.6 revealed the expression of mRNA species, in which an *rev1* encoding exon is joined to an exon containing the 3′ part of the *nef* ORF (**Figure 2**) (Ocwieja et al., 2012). These transcripts are the result of splicing events involving acceptors A4a–c and A8c, and encode a fusion of Rev1 and the C-terminal 80 amino acids of Nef. In reference to its parental proteins Rev and Nef, this fusion protein was named Ref. Although the amount of Ref encoding transcripts exceeded 20% of all completely spliced 1 kb

mRNA species, a fusion protein was hardly detectable. A 12.5 kDa protein representing Ref became only detectable in transfected HEK293T cells treated with the proteasome inhibitor MG132. These findings suggest that the fusion protein is very unstable and are in agreement with the observation that Ref neither exerts Rev activity nor interferes with regular Rev function or HIV-1 particle production (Ocwieja et al., 2012).

EXPRESSION OF A Tat-T FUSION PROTEIN (Vpt) UPON RIBOSOMAL FRAMESHIFTING

In addition to alternative splicing, ribosomal frameshifting represents another mechanism that may result in the expression of fusion proteins. The most prominent example in HIV-1 and related primate lentiviruses is the Gag/Pol polyprotein, which is the result of a −1 frameshift event in the *gag* ORF (**Figure 3**). The frameshift in *pol* depends on a stem-loop structure stalling the translocating ribosome and an upstream heptameric "slippery site" where ribosomal frameshifting occurs (Dinman et al., 2002). While the slippage heptamer (5′-UUUUUUA-3′) itself results in frameshifting, its frequency is increased to about 5% by the adjacent stem-loop structure (Kobayashi et al., 2010; Mouzakis et al., 2013). Interestingly, a similar combination of slippage sequence and RNA secondary structure can be found within the first exon of *tat* (Cohen et al., 1990). This second sequence (5′-UAAAAAG-3′) is highly conserved among HIV-1 strains (Steffy and Wong-Staal, 1991) and has been shown to result in the expression of a cryptic reading frame called T that overlaps with *rev1* and *vpu* (**Figure 3**) (Sonigo et al., 1985; Cohen et al., 1990). Due to a −1 frameshift, this T open reading frame (which does not harbor an initiation codon) is fused to the N-terminus of Tat1, resulting in the expression of a 17 kDa protein called Tat-T or Vpt. Although the frameshift signal is evolutionarily conserved, expression of this fusion protein in primary HIV-1 target cells is questionable. So far, this protein has only been detected upon *in vitro* translation using reticulocyte extract (Cohen et al., 1990). In fact, expression in infected T cells may be prevented by several splice sites disrupting the T open reading frame. In agreement with this, Tat-T/Vpt was not detectable in Jurkat and COS cells transfected with proviral HXBc2 DNA, and 50 different patient sera failed to detect expression of this fusion protein from an expression plasmid (Cohen et al., 1990). Finally, Tat-T does not exert any detectable Tat or Rev activity (Cohen et al., 1990). Nevertheless, even if Tat-T/Vpt is not expressed *in vivo*, it remains to be determined whether or how the frameshift sequence in *tat1* affects translation of regular Tat.

GENE REARRANGEMENTS ENABLE THE EXPRESSION OF A Rev1-Vpu FUSION PROTEIN

In the majority of HIV-1 strains, the *rev1* and *vpu* genes lie in different reading frames and/or are separated by an intervening stop codon (**Figure 4**). However, about 3% of clade A and 20% of clade C viruses as well as some circulating recombinants thereof encode an unusual *rev1-vpu* fusion gene (Kraus et al., 2010). Analysis of primary HIV-1 isolates harboring this ORF revealed that infected PBMCs express a Rev1-Vpu fusion protein of about 14 kDa (Langer et al., 2015). Although prevalence rates may be skewed by sampling biases, it is tempting to speculate that more than 10% of all circulating HIV-1 strains encode this unusual fusion protein, as subtype C viruses are responsible for about 50% of all infections worldwide (Osmanov et al., 2002; Hemelaar et al., 2006, 2011). Cells infected with *rev1-vpu* containing viruses, however, still express regular Vpu at much higher levels than Rev1-Vpu as most *vpu* encoding transcripts lack the initiation codon of *rev1* (Kraus et al., 2010; Ocwieja et al., 2012; Langer et al., 2015): in about 75-90% of all *vpu/env* mRNAs, an intron containing the start codon of Rev1 has been removed due to the usage of splice acceptor A5 (Purcell and Martin, 1993; Ocwieja et al., 2012). Only in 10–25% of the cases, A4 splice acceptors are used and the complete *rev1* ORF is retained. The expression of Rev1-Vpu may be further lowered by leaky scanning, in which the Rev1 initiation codon is skipped due to a weak Kozak sequence.

The characterization of virus pairs differing solely in their ability to express Rev1-Vpu revealed that the presence of this unusual fusion gene does not affect Rev-dependent nuclear export of incompletely spliced viral mRNAs. In agreement with the low Rev1-Vpu:Vpu ratio, downmodulation of CD4, tetherin counteraction and inhibition of NF-κB activation by Vpu were not affected either (Langer et al., 2015). Since the presence of *rev1-vpu* did not enhance viral replication in PBMCs or *ex vivo* infected tonsillar tissue, this gene arrangement does not seem to confer a selection advantage to HIV-1 *per-se*. Interestingly, however, mutations in the *rev1-vpu* intervening region strongly affected Env expression in some viruses (Langer et al., 2015), and the presence of the fusion gene in *rev/vpu/env* expression cassettes impeded pseudotyping of *env*-deficient viruses (Kraus et al., 2010). Previous studies demonstrated that HIV-1 optimizes Env expression throughout the course of infection to increase viral infectivity and transmission while minimizing antibody neutralization and immune activation (Parrish et al., 2013; Krapp et al., 2016). Thus, the expression of Rev1-Vpu may be merely an epiphenomenon of adaptive changes modulating Env expression. This hypothesis is in agreement with the description of several regulatory elements in the *rev1/vpu* region that may modulate leaky scanning and/or induce ribosomal shunting (Schwartz et al., 1990b; Anderson et al., 2007; Krummheuer et al., 2007). Together with the observation that the proportion of *rev1-vpu* encoding viruses does not seem to increase over time (unpublished data), these findings strongly suggest that the fusion gene itself has a neutral phenotype.

SUMMARY AND CONCLUDING REMARKS

The generation of fusion proteins represents a mechanism of increasing the coding potential of viral genomes and has been identified in diverse viruses including primate lentiviruses, foamy, and papilloma viruses (Lambert et al., 1989; Viglianti

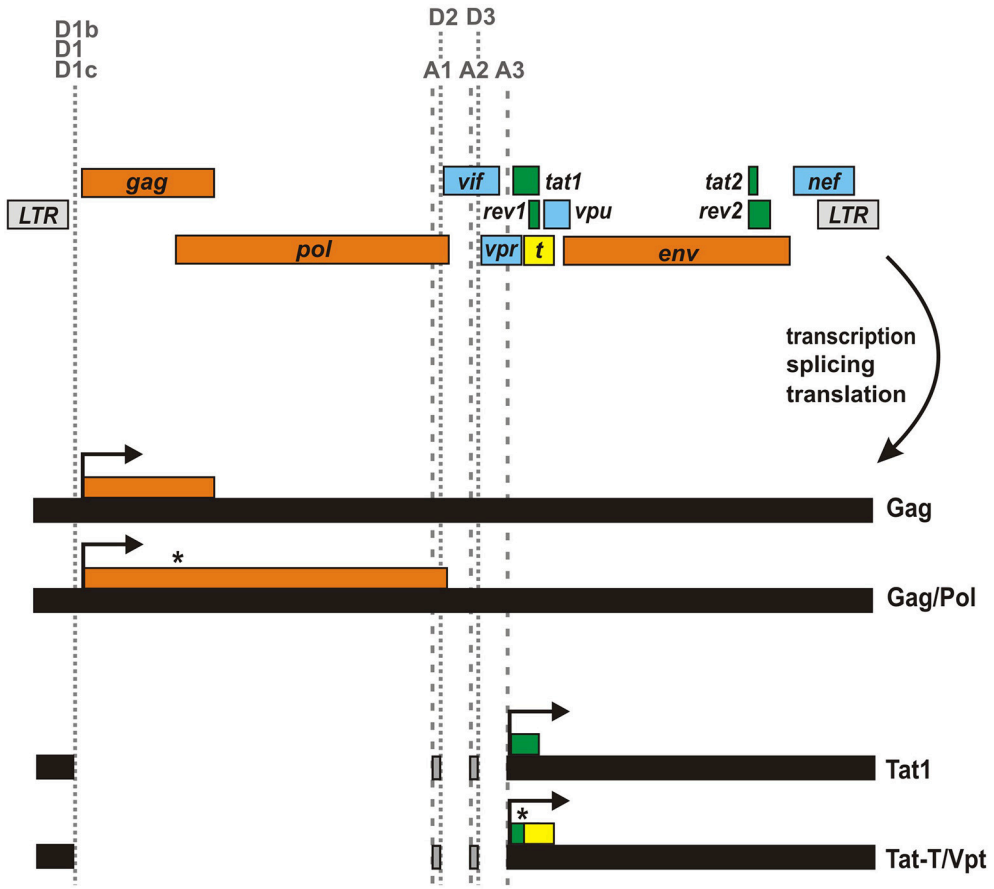


FIGURE 3 | Generation of Gag-Pol and Tat-T/Vpt by ribosomal frameshifting. Unspliced viral mRNA contains a slippery sequence and a downstream RNA secondary structure in *gag* resulting in a frameshift event in about 5% of all translation events. This ribosomal frameshift signal (highlighted by an asterisk) enables the synthesis of the Gag/Pol polyprotein instead of Gag only. A similar frameshift signal in Tat1-encoding mRNA species has been suggested to result in the expression of an unusual protein, in which the N-terminus of Tat1 is fused to the T open reading frame (yellow) encoded by the *tat1/rev1/vpu* locus.

et al., 1990; Lindemann and Rethwilm, 1998). By literally piecing together functional domains of different proteins, viruses may generate fusion products that retain or regulate the activity of their parental proteins and/or even exert entirely novel functions. Although HIV-1 is among the best characterized viruses, relatively little is known about its "fuseome," i.e., the entity of all viral fusion genes and proteins. To date, more than a dozen non-canonical lentiviral fusion proteins have been described (**Table 1**). While two of them, Tat-T/Vpt and Rev1-Vpu, are the result of ribosomal frameshifting and genetic rearrangements, respectively, the remaining ones are expressed due to alternative splicing events. Although for some of them, expression has been confirmed on both mRNA and protein levels, a relevant role for all of these fusion proteins in lentiviral replication remains questionable for several reasons: (1) Most of the fusion proteins were only identified in a few laboratory-adapted viruses. For example, cryptic exon 6D, which is required for the generation of TNV/p28^{tev} and p18^{6Drev}, has only been described for HIV-1 HXB2 and a few closely related viruses (Feinberg et al., 1986; Wright et al., 1986; Benko et al., 1990; Salfeld et al., 1990; Schwartz et al., 1990a; Göttlinger et al., 1992; Neumann et al., 1994; Wentz et al., 1997). Follow-up studies including *in vivo* transcriptome analyses of patient-derived cells failed to detect 6D transcripts in other HIV-1 strains and subtypes suggesting that they might represent an artifact of laboratory-adapted viruses (Furtado et al., 1991; Smith et al., 1992; Purcell and Martin, 1993; Vega et al., 2016). To our knowledge, Rev1-Vpu is the only unusual fusion protein known to be expressed by intact primary isolates of HIV-1 (Langer et al., 2015). (2) The total cellular levels of many fusion proteins are very low. P17^{tev} and Tat-T/Vpt, for example, were detectable upon *in vitro* translation, but not in transfected or infected T cells (Cohen et al., 1990; Furtado et al., 1991). Similarly, Ref was not detectable by Western blotting unless the cells were treated with a proteasome inhibitor (Ocwieja et al., 2012). (3) Although some fusion proteins were shown to exert the activity of their parental proteins, several mutational analyses argue against a crucial role of known fusion proteins in viral replication. Mutation of the splice sites generating exon 6D, for example, revealed that TNV/p28^{tev} is not required for efficient replication of HIV-1

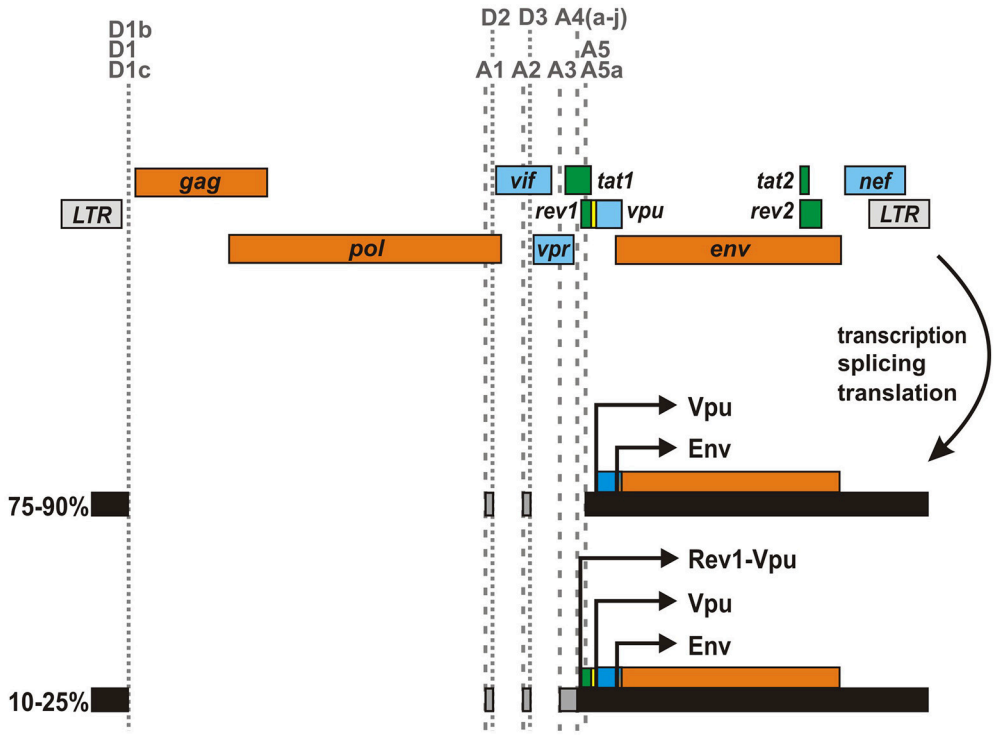


FIGURE 4 | Expression of a *rev1-vpu* fusion gene by certain HIV-1 strains. In some HIV-1 M clade A and C strains, *rev1* and *vpu* are located in the same reading frame and not separated by an intervening stop codon. This unusual gene arrangement results in the expression of a Rev1-Vpu fusion protein from *vpu/env* encoding mRNA species. However, as the majority of *vpu/env* encoding mRNAs lack the *rev1* initiation codon, regular Vpu is expressed at much higher levels than Rev1-Vpu.

in CD4+ T cells (Göttlinger et al., 1992). In fact, elevated usage of exon 6D may even be detrimental for viral replication (Wentz et al., 1997). Similarly, the majority of primary HIV-1 isolates seems to do well without a *rev1-vpu* fusion gene, and gain-of-function mutations did not enhance viral replication in PBMCs or lymphoid tissue (Kraus et al., 2010; Langer et al., 2015).

The observation that fusion proteins are expressed only by a fraction of HIV-1 strains and may be dispensable for viral replication *in vivo* strongly suggests that their expression is just a tolerated epiphenomenon of other adaptive changes. In line with this hypothesis, several studies suggested that cryptic splice sites such as A6 may have evolved to stabilize adjacent suboptimal splice sites and/or increase mRNA stability to balance the ratio of spliced and unspliced HIV-1 transcripts (Lu et al., 1990; Haseltine and Wong-Staal, 1991; Göttlinger et al., 1992; Lützelberger et al., 2006). Furthermore, novel splice sites may also be an (inevitable) result of adaptive changes in regulatory RNA elements, such as shunting structures or Kozak sequences. Mutations generating a *rev1-vpu* fusion gene, for example, have been shown to drastically affect *env* expression (Langer et al., 2015). Finally, fusion proteins may evolve to compensate for detrimental mutations elsewhere in the genome. One striking example has been described by the Berkhout lab, where a Tat-Rev fusion protein evolved to compensate for a mutation of the *rev* initiation codon (Verhoef et al., 2001). The observed Tat-Rev fusion comprised all domains of Rev and allowed the virus to replicate almost as efficiently as the respective wild type control.

No matter whether HIV-1 fusion proteins represent beneficial helpers, neutral factors or even detrimental byproducts, all of them may potentially be immunogenic and serve as T cell epitopes and/or antibody binding sites. To better assess their relevance for viral replication and immune activation, it is therefore crucial to investigate viral mRNA and protein expression in a broad and unbiased manner. Since viral gene expression seems to depend on the cell type and the viral strain (Ocwieja et al., 2012; Vega et al., 2016) rather than the time point of infection (Saltarelli et al., 1996), it is especially important to perform analyses in primary target cells infected with diverse groups and clades of HIV-1. For example, the recent pyrosequencing of CD4+ CD25+ lymphocytes from individuals infected with non-B subtypes revealed that the diversity of splice site usage and the expression of non-canonical transcripts is substantially higher than previously anticipated (Vega et al., 2016). Furthermore, Ocwieja and colleagues hypothesized that cryptic splice donor sites near the 3′ end of the viral RNA such as SD8955 or D6 may also be joined with adjacent exons of the host and result in the expression of chimeric viral-host proteins as previously described for self-inactivating (SIN) retroviral vectors (Almarza et al., 2011; Ocwieja et al., 2012). Remarkably, even defective proviruses that fail to produce infectious viral

particles have recently been shown to express RNA species with unusual exon combinations (Imamichi et al., 2016). Due to large (intron) deletions, these unspliced RNAs may be exported from the nucleus in a Rev/RRE-independent manner, where they are predicted to produce truncated and/or chimeric viral proteins.

Thus, it is very likely that the lentiviral fuseome will further increase, and future analyses will reveal whether some HIV-1 strains express non-canonical fusion proteins with relevant functions *in vivo*.

AUTHOR CONTRIBUTIONS

SL and DS wrote this review article.

ACKNOWLEDGMENTS

We thank Frank Kirchhoff for critical reading of the manuscript and the International Graduate School in Molecular Medicine Ulm for supporting SL. DS was funded by the Deutsche Forschungsgemeinschaft (SPP1923).

REFERENCES

Almarza, D., Bussadori, G., Navarro, M., Mavilio, F., Larcher, F., and Murillas, R. (2011). Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional read-through in keratinocytes transduced with self-inactivating lentiviral vectors. *Gene Ther.* 18, 674–681. doi: 10.1038/gt.2011.12

Anderson, J. L., Johnson, A. T., Howard, J. L., and Purcell, D. F. (2007). Both linear and discontinuous ribosome scanning are used for translation initiation from bicistronic human immunodeficiency virus type 1 env mRNAs. *J. Virol.* 81, 4664–4676. doi: 10.1128/JVI.01028-06

Benko, D. M., Schwartz, S., Pavlakis, G. N., and Felber, B. K. (1990). A novel human immunodeficiency virus type 1 protein, tev, shares sequences with tat, env, and rev proteins. *J. Virol.* 64, 2505–2518.

Carrera, C., Pinilla, M., Pérez-Alvarez, L., and Thomson, M. M. (2010). Identification of unusual and novel HIV type 1 spliced transcripts generated *in vivo*. *AIDS Res. Hum. Retroviruses* 26, 815–820. doi: 10.1089/aid.2010.0011

Cohen, E. A., Lu, Y., Göttlinger, H., Dehni, G., Jalinoos, Y., Sodroski, J. G., et al. (1990). The T open reading frame of human immunodeficiency virus type 1. *J. Acquir. Immune Defic. Syndr.* 3, 601–608.

Cullen, B. R. (1990). The HIV-1 Tat protein: an RNA sequence-specific processivity factor? *Cell* 63, 655–657.

Dinman, J. D., Richter, S., Plant, E. P., Taylor, R. C., Hammell, A. B., and Rana, T. M. (2002). The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5331–5336. doi: 10.1073/pnas.082102199

Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C., and Wong-Staal, F. (1986). HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* 46, 807–817. doi: 10.1016/0092-8674(86)90062-0

Firth, A. E., and Brierley, I. (2012). Non-canonical translation in RNA viruses. *J. Gen. Virol.* 93, 1385–1409. doi: 10.1099/vir.0.042499-0

Furtado, M. R., Balachandran, R., Gupta, P., and Wolinsky, S. M. (1991). Analysis of alternatively spliced human immunodeficiency virus type-1 mRNA species, one of which encodes a novel tat-env fusion protein. *Virology* 185, 258–270. doi: 10.1016/0042-6822(91)90773-5

Göttlinger, H. G., Dorfman, T., Cohen, E. A., and Haseltine, W. A. (1992). The role of the tnv protein and tnv RNA splicing signals in replication of HIV-1 IIIB isolates. *Virology* 189, 618–628. doi: 10.1016/0042-6822(92)90585-D

Haseltine, W. A., and Wong-Staal, F. (1991). *Genetic Structure and Regulation of HIV*. Boca Raton, FL: Raven Press.

Hemelaar, J., Gouws, E., Ghys, P. D., and Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS Lond. Engl.* 20, W13–W23. doi: 10.1097/01.aids.0000247564.73009.bc

Hemelaar, J., Gouws, E., Ghys, P. D., Osmanov, S., and WHO-UNAIDS Network for HIV Isolation and Characterisation (2011). Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS Lond. Engl.* 25, 679–689. doi: 10.1097/QAD.0b013e328342ff93

Imamichi, H., Dewar, R. L., Adelsberger, J. W., Rehm, C. A., O'Doherty, U., Paxinos, E. E., et al. (2016). Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc. Natl. Acad. Sci. U.S.A.* 113, 8783–8788. doi: 10.1073/pnas.1609057113

Kobayashi, Y., Zhuang, J., Peltz, S., and Dougherty, J. (2010). Identification of a cellular factor that modulates HIV-1 programmed ribosomal frameshifting. *J. Biol. Chem.* 285, 19776–19784. doi: 10.1074/jbc.M109.085621

Krapp, C., Hotter, D., Gawanbacht, A., McLaren, P. J., Kluge, S. F., Stürzel, C. M., et al. (2016). Guanylate Binding Protein (GBP) 5 is an interferon-inducible inhibitor of HIV-1 infectivity. *Cell Host Microbe* 19, 504–514. doi: 10.1016/j.chom.2016.02.019

Kraus, M. H., Parrish, N. F., Shaw, K. S., Decker, J. M., Keele, B. F., Salazar-Gonzalez, J. F., et al. (2010). A rev1-vpu polymorphism unique to HIV-1 subtype A and C strains impairs envelope glycoprotein expression from rev-vpu-env cassettes and reduces virion infectivity in pseudotyping assays. *Virology* 397, 346–357. doi: 10.1016/j.virol.2009.11.019

Krummheuer, J., Johnson, A. T., Hauber, I., Kammler, S., Anderson, J. L., Hauber, J., et al. (2007). A minimal uORF within the HIV-1 vpu leader allows efficient translation initiation at the downstream env AUG. *Virology* 363, 261–271. doi: 10.1016/j.virol.2007.01.022

Lambert, P. F., Hubbert, N. L., Howley, P. M., and Schiller, J. T. (1989). Genetic assignment of multiple E2 gene products in bovine papillomavirus-transformed cells. *J. Virol.* 63, 3151–3154.

Langer, S. M., Hopfensperger, K., Iyer, S. S., Kreider, E. F., Learn, G. H., Lee, L.-H., et al. (2015). A naturally occurring rev1-vpu fusion gene does not confer a fitness advantage to HIV-1. *PLoS ONE* 10:e0142118. doi: 10.1371/journal.pone.0142118

Lindemann, D., and Rethwilm, A. (1998). Characterization of a human foamy virus 170-kilodalton Env-Bet fusion protein generated by alternative splicing. *J. Virol.* 72, 4088–4094.

Lützelberger, M., Reinert, L. S., Das, A. T., Berkhout, B., and Kjems, J. (2006). A novel splice donor site in the gag-pol gene is required for HIV-1 RNA stability. *J. Biol. Chem.* 281, 18644–18651. doi: 10.1074/jbc.M513698200

Lu, X. B., Heimer, J., Rekosh, D., and Hammarskjöld, M. L. (1990). U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc. Natl. Acad. Sci. U.S.A.* 87, 7598–7602. doi: 10.1073/pnas.87.19.7598

Malim, M. H., Böhnlein, S., Hauber, J., and Cullen, B. R. (1989). Functional dissection of the HIV-1 Rev trans-activator–derivation of a trans-dominant repressor of Rev function. *Cell* 58, 205–214. doi: 10.1016/0092-8674(89)90416-9

Mouzakis, K. D., Lang, A. L., Vander Meulen, K. A., Easterday, P. D., and Butcher, S. E. (2013). HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res.* 41, 1901–1913. doi: 10.1093/nar/gks1254

Neumann, M., Harrison, J., Saltarelli, M., Hadziyannis, E., Erfle, V., Felber, B. K., et al. (1994). Splicing variability in HIV type 1 revealed by quantitative RNA polymerase chain reaction. *AIDS Res. Hum. Retroviruses* 10, 1531–1542. doi: 10.1089/aid.1994.10.1531

Ocwieja, K. E., Sherrill-Mix, S., Mukherjee, R., Custers-Allen, R., David, P., Brown, M., et al. (2012). Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res.* 40, 10345–10355. doi: 10.1093/nar/gks753

Osmanov, S., Pattou, C., Walker, N., Schwardländer, B., Esparza, J., and WHO-UNAIDS Network for HIV Isolation and Characterization (2002).

Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J. Acquir. Immune Defic. Syndr.* 29, 184–190. doi: 10.1097/00042560-200202010-00013

Parrish, N. F., Gao, F., Li, H., Giorgi, E. E., Barbian, H. J., Parrish, E. H., et al. (2013). Phenotypic properties of transmitted founder HIV-1. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6626–6633. doi: 10.1073/pnas.1304288110

Pearson, L., Garcia, J., Wu, F., Modesti, N., Nelson, J., and Gaynor, R. (1990). A transdominant tat mutant that inhibits tat-induced gene expression from the human immunodeficiency virus long terminal repeat. *Proc. Natl. Acad. Sci. U.S.A.* 87, 5079–5083. doi: 10.1073/pnas.87.13.5079

Pereira, E. A., and daSilva, L. L. (2016). HIV-1 Nef: taking control of protein trafficking. *Traffic Cph. Den.* 17, 976–996. doi: 10.1111/tra.12412

Purcell, D. F., and Martin, M. A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol.* 67, 6365–6378.

Salfeld, J., Göttlinger, H. G., Sia, R. A., Park, R. E., Sodroski, J. G., and Haseltine, W. A. (1990). A tripartite HIV-1 tat-env-rev fusion protein. *EMBO J.* 9, 965–970.

Saltarelli, M. J., Hadziyannis, E., Hart, C. E., Harrison, J. V., Felber, B. K., Spira, T. J., et al. (1996). Analysis of human immunodeficiency virus type 1 mRNA splicing patterns during disease progression in peripheral blood mononuclear cells from infected individuals. *AIDS Res. Hum. Retroviruses* 12, 1443–1456. doi: 10.1089/aid.1996.12.1443

Schwartz, S., Felber, B. K., Benko, D. M., Fenyö, E. M., and Pavlakis, G. N. (1990a). Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J. Virol.* 64, 2519–2529.

Schwartz, S., Felber, B. K., Fenyö, E. M., and Pavlakis, G. N. (1990b). Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J. Virol.* 64, 5448–5456.

Smith, J., Azad, A., and Deacon, N. (1992). Identification of two novel human immunodeficiency virus type 1 splice acceptor sites in infected T cell lines. *J. Gen. Virol.* 73(Pt 7), 1825–1828. doi: 10.1099/0022-1317-73-7-1825

Sodroski, J., Rosen, C., Wong-Staal, F., Salahuddin, S. Z., Popovic, M., Arya, S., et al. (1985). Trans-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. *Science* 227, 171–173. doi: 10.1126/science.2981427

Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., et al. (1985). Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. *Cell* 42, 369–382. doi: 10.1016/S0092-8674(85)80132-X

Steffy, K., and Wong-Staal, F. (1991). Genetic regulation of human immunodeficiency virus. *Microbiol. Rev.* 55, 193–205.

Vega, Y., Delgado, E., de la Barrera, J., Carrera, C., Zaballos, Á., Cuesta, I., et al. (2016). Sequence analysis of *in vivo*-expressed HIV-1 spliced RNAs reveals the usage of new and unusual splice sites by viruses of different subtypes. *PLoS ONE* 11:e0158525. doi: 10.1371/journal.pone.0158525

Verhoef, K., Bilodeau, P. S., van Wamel, J. L., Kjems, J., Stoltzfus, C. M., and Berkhout, B. (2001). Repair of a Rev-minus human immunodeficiency virus type 1 mutant by activation of a cryptic splice site. *J. Virol.* 75, 3495–3500. doi: 10.1128/JVI.75.7.3495-3500.2001

Viglianti, G. A., Sharma, P. L., and Mullins, J. I. (1990). Simian immunodeficiency virus displays complex patterns of RNA splicing. *J. Virol.* 64, 4207–4216.

Vives, E., Charneau, P., van Rietschoten, J., Rochat, H., and Bahraoui, E. (1994). Effects of the Tat basic domain on human immunodeficiency virus type 1 transactivation, using chemically synthesized Tat protein and Tat peptides. *J. Virol.* 68, 3343–3353.

Wentz, M. P., Moore, B. E., Cloyd, M. W., Berget, S. M., and Donehower, L. A. (1997). A naturally arising mutation of a potential silencer of exon splicing in human immunodeficiency virus type 1 induces dominant aberrant splicing and arrests virus production. *J. Virol.* 71, 8542–8551.

Wright, C. M., Felber, B. K., Paskalis, H., and Pavlakis, G. N. (1986). Expression and characterization of the trans-activator of HTLV-III/LAV virus. *Science* 234, 988–992. doi: 10.1126/science.3490693

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Langer and Sauter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

MINI REVIEW
published: 16 March 2017
doi: 10.3389/fmicb.2017.00390

Driving HIV-1 into a Vulnerable Corner by Taking Advantage of Viral Adaptation and Evolution

*Shigeyoshi Harada and Kazuhisa Yoshimura**

AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan

Anti-retroviral therapy (ART) is crucial for controlling human immunodeficiency virus type-1 (HIV-1) infection. Recently, progress in identifying and characterizing highly potent broadly neutralizing antibodies has provided valuable templates for HIV-1 therapy and vaccine design. Nevertheless, HIV-1, like many RNA viruses, exhibits genetically diverse populations known as quasispecies. Evolution of quasispecies can occur rapidly in response to selective pressures, such as that exerted by ART and the immune system. Hence, rapid viral evolution leading to drug resistance and/or immune evasion is a significant barrier to the development of effective HIV-1 treatments and vaccines. Here, we describe our recent investigations into evolutionary pressure exerted by anti-retroviral drugs and monoclonal neutralizing antibodies (NAbs) on HIV-1 envelope sequences. We also discuss sensitivities of HIV-1 escape mutants to maraviroc, a CCR5 inhibitor, and HIV-1 sensitized to NAbs by small-molecule CD4-mimetic compounds. These studies help to develop an understanding of viral evolution and escape from both anti-retroviral drugs and the immune system, and also provide fundamental insights into the combined use of NAbs and entry inhibitors. These findings of the adaptation and evolution of HIV in response to drug and immune pressure will inform the development of more effective antiviral therapeutic strategies.

Keywords: HIV-1, antiretroviral therapy, neutralizing antibody, evolution, escape

OPEN ACCESS

Edited by:
Akio Adachi,
University of Tokushima, Japan

Reviewed by:
Takamasa Ueno,
Kumamoto University, Japan
Thorsten Demberg,
Immatics Biotechnologies, Germany
Keisuke Yusa,
National Institute of Health Sciences, Japan

****Correspondence:***
Kazuhisa Yoshimura
ykazu@nih.go.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *31 October 2016*
Accepted: *24 February 2017*
Published: *16 March 2017*

Citation:
Harada S and Yoshimura K (2017) Driving HIV-1 into a Vulnerable Corner by Taking Advantage of Viral Adaptation and Evolution. Front. Microbiol. 8:390. doi: 10.3389/fmicb.2017.00390

INTRODUCTION

Human immunodeficiency virus type-1 (HIV-1) exhibits extremely high genetic diversity (Rambaut et al., 2004) indicating that rapidly changing genetic variation can confer on the virus the capacity to escape the immune system and anti-retroviral therapy (ART). The HIV-1 components presenting the highest degree of sequence diversity are the surface-expressed viral envelope glycoproteins (Env), which are prime targets for both entry inhibitors and neutralizing antibodies (NAbs) (Goulder and Watkins, 2008).

The function of Env is to facilitate the entry of HIV-1 into the target cell, a process mediated by recognition of the CD4 receptor and coreceptor (usually CCR5 or CXCR4) on the cellular membrane (Dalgleish et al., 1984; Klatzmann et al., 1984; Choe et al., 1996; Deng et al., 1996; Doranz et al., 1996; Dragic et al., 1996; Feng et al., 1996). Env is composed of the surface glycoprotein, gp120, and the transmembrane glycoprotein, gp41, which associate as a non-covalent complex to form a single subunit of a trimeric viral envelope spike (Wyatt and Sodroski, 1998). Gp120 is responsible for interactions with CD4 and the coreceptor, whereas gp41 anchors the Env

machinery at the viral membrane and induces membrane fusion during viral entry (Freed and Martin, 1995; Bazan et al., 1998).

Many entry inhibitors have been developed to block the interaction of Env with the CD4 receptor, the coreceptor, or the fusion reaction. Currently, two entry inhibitors have been approved for clinical use, the fusion inhibitor, enfuvirtide (T-20) (Robertson, 2003), and the CCR5 inhibitor, maraviroc (MVC) (Dorr et al., 2005; Gulick et al., 2008). As with any anti-retroviral drug, HIV can develop resistance to T-20 and MVC. The major mechanism of resistance to T-20 is caused by mutations within the binding site on the HR1 region of gp41 (Greenberg and Cammack, 2004) (**Figure 1C**). On the other hand, clinical resistance to MVC involves different genetic alterations in *env* giving rise to highly divergent Env phenotypes (Roche et al., 2013). Potential molecular mechanisms of resistance to MVC include tropism switching to CXCR4-using (X4) viruses (Westby et al., 2006; Raymond et al., 2015), increased kinetics of the entry step (Reeves et al., 2002; Putcharoen et al., 2012), increased affinity for CD4 and/or CCR5 (Agrawal-Gamse et al., 2009; Pugach et al., 2009; Pfaff et al., 2010; Ratcliff et al., 2013), and utilization of MVC-bound CCR5 for entry (Pugach et al., 2007; Westby et al., 2007; Tilton et al., 2010; Roche et al., 2011).

In recent years, progress in identifying and characterizing highly potent broadly NAbs (bNAbs), has provided valuable templates for HIV-1 therapy and vaccine design (Kwong and Mascola, 2012; Kwong et al., 2013; Burton and Mascola, 2015; Burton and Hangartner, 2016). However, attempts to elicit such highly potent bNAbs by immunization have not been successful, due in part to the high genetic diversity of Env and the complex escape mechanisms employed by Env (Seaman et al., 2010).

Moreover, the replication capacity of HIV-1 is largely related to the efficiency of viral entry (Arts and Quinones-Mateu, 2003; Rangel et al., 2003). In this respect, evolutionary patterns of Env are important, and selective pressures exerted by NAbs and anti-retroviral drugs can contribute to its evolution. Thus, elucidation of these patterns would inform the development of more effective antiviral therapeutic strategies.

Recently, we investigated dynamic features of selective pressure on Env by assessing NAb sensitivities of HIV-1 escape mutants from MVC, and small-molecule CD4-mimetic compounds (CD4mc) that sensitize HIV-1 to NAbs. Thus, we summarize these recent advances and discuss the application of these findings to the development of more effective combinations of NAbs and anti-retroviral drugs.

FUNDAMENTALS OF HIV ENTRY

Entry of HIV-1 into a target cell involves interactions between Env and the two-receptor mechanism involving CD4 and the coreceptor. This interaction activates conformational changes in Env that lead to the membrane fusion reaction (Sattentau and Moore, 1995) (**Figure 1B**).

Gp120 is composed of five conserved regions (C1 to C5) that are interspersed with five variable regions (V1 to V5) (Starcich et al., 1986) (**Figure 1C**). The CD4 binding site (CD4bs) and especially the Phe 43 cavity, where Phe 43 of CD4 contacts gp120, are highly conserved among the different subtypes (Kwong et al., 1998). Following the binding of CD4 and gp120, the gp120 core undergoes conformational changes, moving from a rigid (unliganded) to a flexible state, allowing a subsequent interaction with the coreceptor (Myszka et al., 2000) (**Figure 1B**). Binding of gp120 to the coreceptor triggers further conformational changes in Env that fuse the viral membrane with the target cell membrane (Chan and Kim, 1998). Current models suggest the V3 tip interacts with the coreceptor second extracellular loop (ECL2), whereas the gp120 bridging sheet and the V3 stem interact with the coreceptor N terminus (Brelot et al., 1999; Farzan et al., 1999; Cormier and Dragic, 2002; Huang et al., 2005) (**Figure 1A**).

PRESSURE OF NAbs ON THE EVOLUTION OF Env

Recently, bNAbs have been isolated from HIV-1-infected individuals. Most major target specificities of these bNAbs have been mapped to various sites on Env, and include the V2 N160 glycan (V2 apex), the V3 N332 glycan (high-mannose patch), the CD4bs, the gp120/41 interface region, the fusion peptide (FP), and the membrane proximal external region (MPER) of gp41 (Burton and Mascola, 2015; Burton and Hangartner, 2016; Kong et al., 2016; van Gils et al., 2016). In addition, CD4 binding exposes highly conserved cryptic epitopes recognized by V3-directed or CD4-induced (CD4i) NAbs, which recognize the coreceptor binding site (Kwong and Mascola, 2012) (**Figure 1B**).

The V3-directed NAb, KD-247, is a humanized NAb with potent neutralizing activity. The epitope recognized by KD-247 was mapped to the IGPGR sequence of the V3-tip, which covers about half of subtype B. A phase-1b clinical study indicates that KD-247 reduces viral load in patients with chronic HIV-1 infection (Matsushita et al., 2015). However, HIV-1 can escape from the adaptive immune responses, and can become resistant to all anti-retroviral drugs. Therefore, in our previous *in vitro* study, we induced resistant variants against KD-247 using the JR-FL strain (Yoshimura et al., 2006). Resistance against KD-247 was associated with G314E substitution in the epitope on the V3-tip. Unexpectedly, the KD-247-resistant variant exhibited higher sensitivity to CCR5 inhibitors (TAK-779, aplaviroc and SCH-C) compared with the parental virus. Furthermore, our data showed strong synergistic interactions between KD-247 and CCR5 inhibitors (Yoshimura et al., 2006).

In addition to our studies, recent investigations of passive NAb therapy in HIV-infected individuals demonstrated that particular bNAbs could reduce levels of plasma viremia and suppress neutralization-sensitive viruses (Caskey et al., 2015; Lynch et al., 2015; Matsushita et al., 2015; Bar et al., 2016). However, a single use of NAbs could not suppress HIV completely and poses the danger of inducing escape variants *in vivo*. These findings suggest that combination strategies containing NAbs are needed to maintain virus suppression and prevent appearance of NAb-escape variants. Therefore, in the near future, combinations of NAbs and CCR5 inhibitors are likely to be efficient weapons against HIV-1.

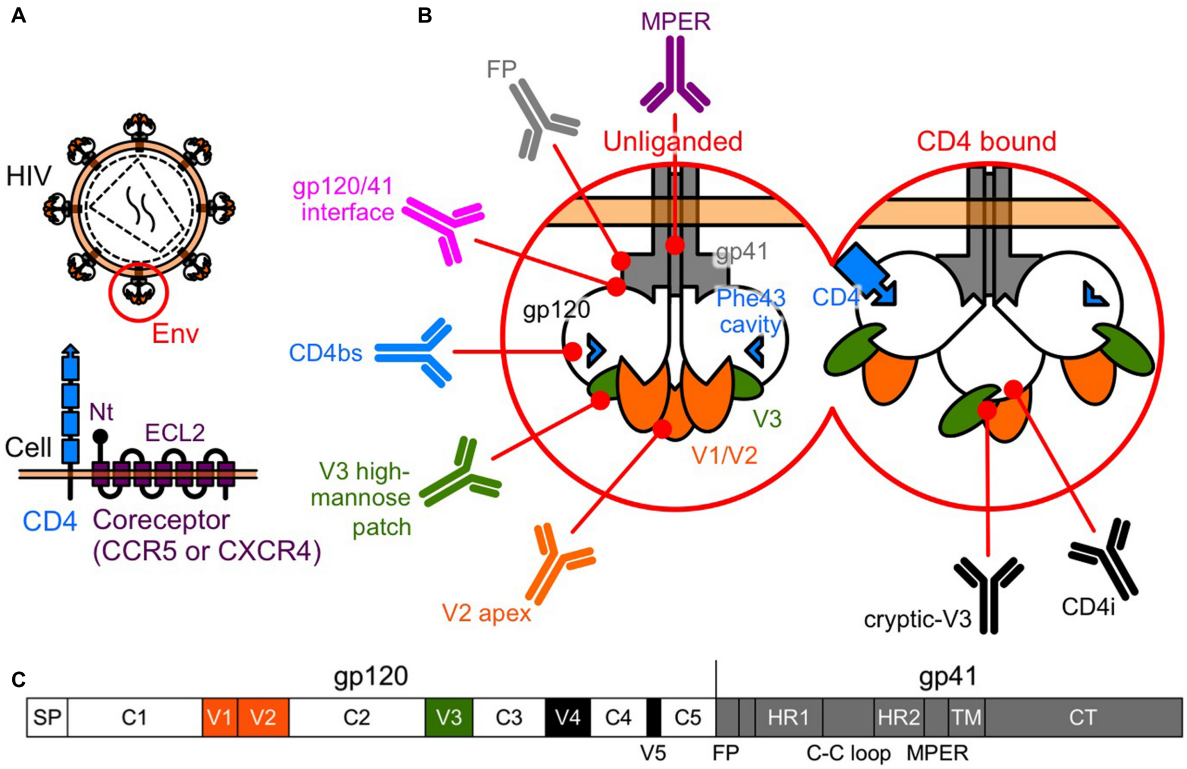


FIGURE 1 | Human immunodeficiency virus type-1 (HIV-1) Env. (A) Entry of HIV-1 into a host cell involves interactions between the Env and the two-receptor mechanism of CD4 and the coreceptor. **(B)** Tertiary schematic view of HIV-1 Env. Following the binding of CD4 and gp120, gp120 undergoes conformational changes, moving from a rigid (unliganded) to a flexible state, allowing a subsequent interaction with the coreceptors. bNAbs have been identified that target the V2 apex, the V3 high-mannose patch, the CD4bs, the gp120/41 interface, the FP, and the MPER of gp41. In the CD4-bound state, a larger area is uncovered and potentially available for recognition by NAbs, such as V3-directed or CD4i, which recognize the conserved coreceptor-binding site. **(C)** Linear schematic view of HIV-1 Env. Gp120 is composed of five conserved regions (C1 to C5) that are interspersed with five variable regions (V1 to V5).

EFFECT OF MVC-RESISTANCE MUTATIONS ON SENSITIVITY TO NAbs

The main mechanism of resistance to MVC appears to be related to changes in the V3 region, which enables the virus to utilize MVC-bound CCR5 coreceptors. The resistance is characterized by reductions in the maximal percent inhibition (MPI) value rather than shifts in the IC_{50} value (Pugach et al., 2007; Roche et al., 2011).

Pugach et al. noted that resistant variants against two CCR5 inhibitors (vicriviroc and AD101) were more sensitive to several types of NAbs compared with the parental virus (Pugach et al., 2008; Berro et al., 2009). Subsequently, we have reported the resistance induction of the primary KP-5P virus (subtype B, R5) against MVC *in vitro* (Yoshimura et al., 2014). Resistance to MVC was associated with V200I, T297I, K305R, and M434I substitutions near the CCR5 binding site. This MVC-resistant variant also exhibited extremely high sensitivity to three NAbs: b12 (CD4bs), 4E9C (CD4i), and KD-247. These results indicated that the MVC-resistance mutations might improve the accessibility of epitopes for the NAbs and, therefore, be incompatible with resistance to the NAbs (Yoshimura et al., 2014). More recently, Kuwata et al. (2015) showed that resistant variants against a CCR5 inhibitor, cenicriviroc, also became sensitive to three NAbs: VRC01 (CD4bs), 4E9C, and 0.5γ (V3).

Another mechanism of resistance to MVC appears to be by a change in coreceptor tropism from CCR5 to CXCR4, or by the selection of minority variants of X4 or dual/mixed viruses (Westby et al., 2006). Indeed, Raymond et al. (2015) has reported that half of MVC-treated patients who experienced virological failure harbored X4 viruses at failure. Remarkably, previous studies have shown that early X4 variants are more sensitive to NAbs compared with their coexisting R5 variants (Ganesh et al., 2004; Lusso et al., 2005; Margolis and Shattock, 2006; Bunnik et al., 2007). In addition, increased CCR5 affinity is also a potential resistance mechanism, but we have shown that low-CCR5 affinity-adapted variants also became sensitive to CD4bs and CD4i NAbs (Yoshimura et al., 2014). Thus, several studies have demonstrated diverse resistance mechanisms against MVC, but all these resistance pathways might drive viral evolution into a corner, escape from which would require high sensitivity to NAbs. Moreover, these observations indicate that MVC and NAbs might limit the emergence of mutants that are resistant to each other, supporting the clinical use of combination therapy (**Figure 2A**).

However, it is not clear whether patients' plasma IgG under MVC treatment can induce mutations in Env to

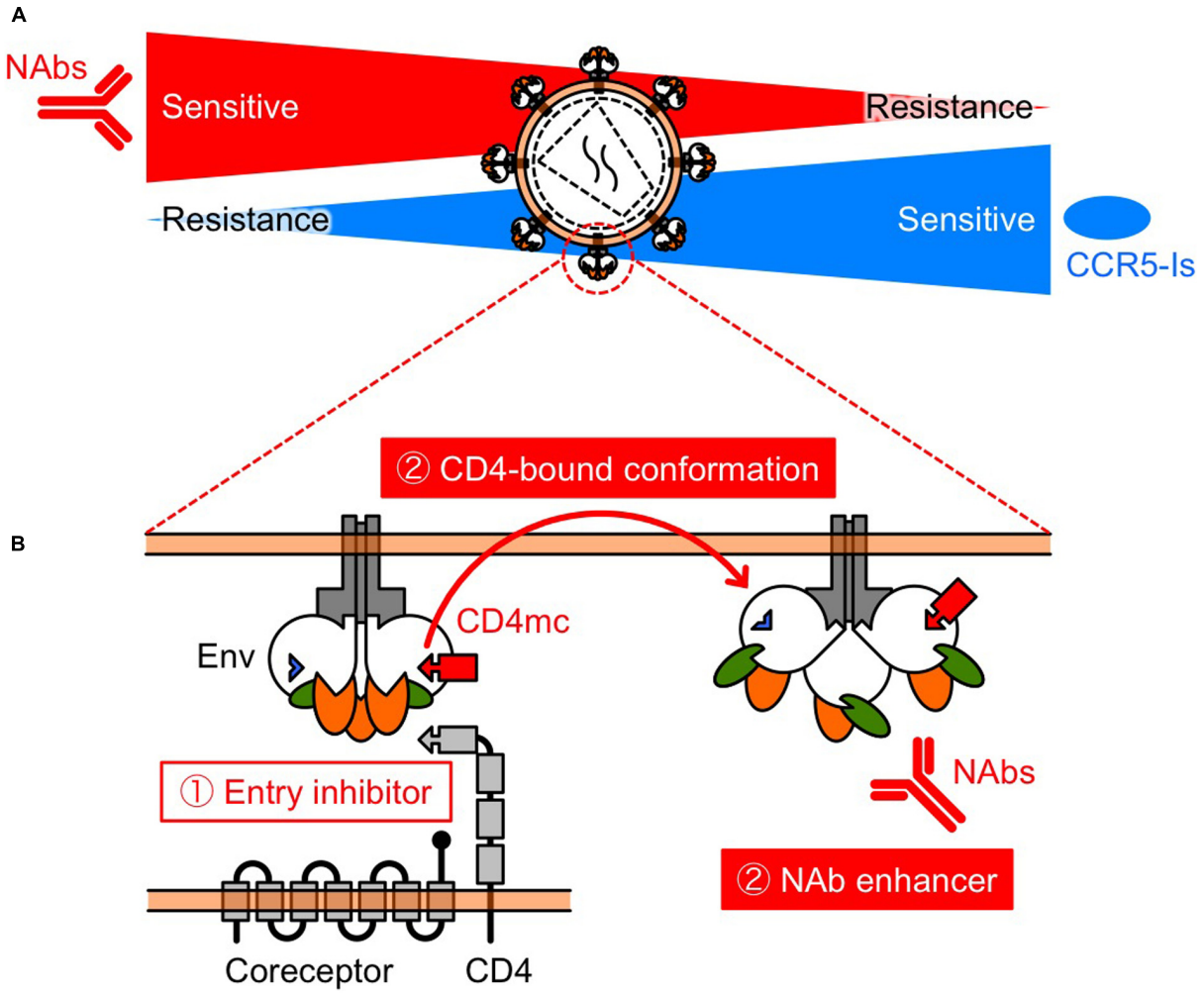


FIGURE 2 | Schematic view of the relationship between NAb resistance and CCR5 inhibitor resistance (A), and the function of CD4mc (B). **(A)** Selection pressures on HIV-1 Env by NAbs and/or CCR5 inhibitors might turn the tide in the fight against HIV-1. The relationship indicates that NAbs and CCR5 inhibitors may restrict the emergence of variants that are resistant to each other. **(B)** CD4mc is a bifunctional entry inhibitor. The use of bifunctional entry inhibitors that display direct blockade of viral entry and exposure of epitopes to NAbs, should be effective in passive NAb immunization.

enhance neutralizing activity. We are currently investigating the relationship between NAb responses and MVC treatment using patients' plasma IgGs before and after MVC-containing combination ART (cART). Moreover, we think treatment with particular entry inhibitors and/or CD4mc can induce bNAbs *in vivo*; however, we await results for this prediction. Thus, we will perform experiments in animal models to induce or enhance NAbs using novel entry inhibitors and/or CD4mc treatment.

ANTI-RETROVIRAL PRESSURE ON THE SELECTION OF Env

Evolution of HIV-1 helps it to evade NAbs (Moore et al., 2012; Liao et al., 2013; Bouvin-Pley et al., 2014). cART, however, results in a reduction in the virus population size, which creates a genetic bottleneck. *In vivo* studies indicate that the bottleneck affects not only drug-target regions (e.g., reverse transcriptase), but also other regions of the viral genome, including the Env region (Zhang et al., 1994; Sheehy et al., 1996; Delwart et al., 1998; Nijhuis et al., 1998; Ibanez et al., 2000; Kitrinos et al., 2005; Charpentier et al., 2006; Nora et al., 2007). The population dynamics of the Env region might be important when bNAbs and novel entry inhibitors become available in the near future. However, it is hard to observe effects of an anti-retroviral drug-induced bottleneck on the Env region *in vivo*.

Thus, we induced variants against anti-retroviral drugs using primary swarm isolates (Harada et al., 2013). As a result, the phylogenetic clustering of raltegravir (an integrase inhibitor)-, lamivudine (a reverse transcriptase inhibitor)- and saquinavir (a protease inhibitor)-induced variants was entirely distinct from that of non-drug-treated controls. Among these drug-induced variants, the variable regions of gp120 were very similar to each other. Conversely, the non-drug-treated variant was quite

different from the drug-induced variants. These results imply that, under selective pressure of non-entry inhibitors, the virus may choose a representative Env sequence from the viral population to gain a growth advantage (Harada et al., 2013). In addition to our results, a supporting study by Mesplede et al. (2015) showed that treatment with dolutegravir (an integrase inhibitor) results in a reduction in viral genetic diversity. Further studies are needed to confirm our observations, but these results may provide a new paradigm for viral evolution in the novel NAb plus anti-retroviral drug combination therapy era.

CD4mc CAN EXPOSE HIV-1 NEUTRALIZATION EPITOPES

Binding of CD4 to gp120, is the first essential step of the entry process. The multiple contacts made by Phe 43 and Arg 59 of CD4 with gp120 residues in CD4bs contribute significantly to CD4–gp120 binding (Kwong et al., 1998). The critical Phe 43 of CD4 becomes buried in a binding pocket of gp120, termed the Phe 43 cavity. This cavity is known to be highly conserved among the different subtypes and is therefore considered a particularly interesting target for inhibitors of CD4–gp120 interaction (Kwong et al., 1998).

Molecules that mimic the CD4 receptor, such as soluble CD4 (sCD4), CD4 immunoadhesin (CD4-Ig), sCD4 mini-proteins, and CD4mc have been developed (Vita et al., 1999; Grupping et al., 2012). sCD4, CD4-Ig, and sCD4 mini-protein have been studied as potential therapeutics (Smith et al., 1987; Fisher et al., 1988; Hussey et al., 1988; Jacobson et al., 2000; Fletcher et al., 2007; Dereuddre-Bosquet et al., 2012; Gardner et al., 2015). These studies in patients and non-human primate models have provided proof of principle that viral entry can be successfully blocked *in vivo*. In particular, Gardner et al. (2015) demonstrated that eCD4-Ig, a fusion of CD4-Ig with a small CCR5-mimetic peptide, was on average more potent, and much broader than bNAbs. Moreover, adeno-associated virus-delivered eCD4-Ig, provided durable protection for immunized monkeys against high-dose intravenous SHIV challenge (Gardner et al., 2015).

The prototype of CD4mc, NBD-556, was identified in a screen for inhibitors of the CD4–gp120 interaction (Zhao et al., 2005). We and others have been exploring the potential of NBD-556-derived CD4mc as a novel class of HIV entry inhibitor (Madani et al., 2004, 2008, 2014, 2016, 2017; Narumi et al., 2010, 2011, 2013; Yamada et al., 2010; Yoshimura et al., 2010; Lalonde et al., 2011, 2012, 2013; Courter et al., 2014; Richard et al., 2015; Melillo et al., 2016; Mizuguchi et al., 2016; Ohashi et al., 2016). The binding of CD4mc in the Phe 43 cavity blocks CD4-gp120 interaction and, induces conformational changes in gp120 similar to those observed upon sCD4 binding (**Figure 2B**) (Schon et al., 2006; Haim et al., 2009; Curreli et al., 2014; Kwon et al., 2014). sCD4 significantly enhance neutralization by CD4i (Thali et al., 1993) and some V3 NAbs (Lusso et al., 2005). Remarkably, CD4i and V3 NAbs are present in HIV-infected individuals during the early stage of infection (Decker et al., 2005). Consequently, we hypothesized that CD4mc can cause exposure of cryptic epitopes to antibodies, allowing virus neutralization. As a result, combinations of CD4mc (NBD-556 or YYA-021) with CD4i or V3 NAbs produced strong synergistic antiviral interactions (Yamada et al., 2010; Yoshimura et al., 2010) (**Figures 1**, **2B**). Moreover, we found that CD4mc sensitized a clinical isolate to autologous plasma antibodies from the same time point (Yoshimura et al., 2010).

Recently, this approach has been extended to combining vaccine with CD4mc. In studies using prototypic CD4mc BNM compounds, Madani et al. (2014, 2016) demonstrated that CD4mc sensitized the virus to antibodies elicited by immunization of humans and monkeys. These studies establish the proof of concept that CD4mc can sensitize primary viruses to antibodies that are present in plasma of infected or vaccinated individuals. In addition, Richard et al. (2015) reported that CD4mc could efficiently sensitize primary CD4 T cells from HIV-1-infected individuals to antibody-dependent cell-mediated cytotoxicity (ADCC) mediated by autologous sera and effector cells.

Based on these results, further studies are needed to investigate the effectiveness of delivery methods of CD4mc. Small molecules such as CD4mc have many advantages over conventional immunotherapeutic agents, including ease of production and the potential for oral administration. Furthermore, the use of bifunctional entry inhibitors that display direct blockade of viral entry and exposure of epitopes to NAbs should be effective in passive NAb immunization.

CONCLUSION

Extensive genetic diversity in the Env region presents significant obstructions to the development of promising therapies and vaccines against HIV-1. However, selection pressures on the Env region by NAbs, entry inhibitors, and/or non-entry antiviral inhibitors, might turn the tide in the fight against HIV-1. Moreover, bifunctional entry inhibitors such as CD4mc might potentiate these selection pressures. Thus, by taking advantage of the adaptation and evolution of HIV resulting from drug and immune pressure, we might drive HIV-1 into a vulnerable corner.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the Ministry of Education, Culture, Sports, Science and Technology (JSPS KAKENHI Grant Number 15K08125), and Japan Agency for Medical Research and Development (AMED).

REFERENCES

Agrawal-Gamse, C., Lee, F. H., Haggarty, B., Jordan, A. P., Yi, Y., Lee, B., et al. (2009). Adaptive mutations in a human immunodeficiency virus type 1 envelope protein with a truncated V3 loop restore function by improving interactions with CD4. *J. Virol.* 83, 11005–11015. doi: 10.1128/JVI.01238-09

Arts, E. J., and Quinones-Mateu, M. E. (2003). Sorting out the complexities of HIV-1 fitness. *AIDS* 17, 780–781. doi: 10.1097/00002030-200303280-00026

Bar, K. J., Sneller, M. C., Harrison, L. J., Justement, J. S., Overton, E. T., Petrone, M. E., et al. (2016). Effect of HIV antibody VRC01 on viral rebound after treatment interruption. *N. Engl. J. Med.* 375, 2037–2050. doi: 10.1056/NEJMoa1608243

Bazan, H. A., Alkhatib, G., Broder, C. C., and Berger, E. A. (1998). Patterns of CCR5, CXCR4, and CCR3 usage by envelope glycoproteins from human immunodeficiency virus type 1 primary isolates. *J. Virol.* 72, 4485–4491.

Berro, R., Sanders, R. W., Lu, M., Klasse, P. J., and Moore, J. P. (2009). Two HIV-1 variants resistant to small molecule CCR5 inhibitors differ in how they use CCR5 for entry. *PLoS Pathog.* 5:e1000548. doi: 10.1371/journal.ppat.1000548

Bouvin-Pley, M., Morgand, M., Meyer, L., Goujard, C., Moreau, A., Mouquet, H., et al. (2014). Drift of the HIV-1 envelope glycoprotein gp120 toward increased neutralization resistance over the course of the epidemic: a comprehensive study using the most potent and broadly neutralizing monoclonal antibodies. *J. Virol.* 88, 13910–13917. doi: 10.1128/JVI.02083-14

Brelot, A., Heveker, N., Adema, K., Hosie, M. J., Willett, B., and Alizon, M. (1999). Effect of mutations in the second extracellular loop of CXCR4 on its utilization by human and feline immunodeficiency viruses. *J. Virol.* 73, 2576–2586.

Bunnik, E. M., Quakkelaar, E. D., Van Nuenen, A. C., Boeser-Nunnink, B., and Schuitemaker, H. (2007). Increased neutralization sensitivity of recently emerged CXCR4-using human immunodeficiency virus type 1 strains compared to coexisting CCR5-using variants from the same patient. *J. Virol.* 81, 525–531. doi: 10.1128/JVI.01983-06

Burton, D. R., and Hangartner, L. (2016). Broadly neutralizing antibodies to HIV and their role in vaccine design. *Annu. Rev. Immunol.* 34, 635–659. doi: 10.1146/annurev-immunol-041015-055515

Burton, D. R., and Mascola, J. R. (2015). Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat. Immunol.* 16, 571–576. doi: 10.1038/ni.3158

Caskey, M., Klein, F., Lorenzi, J. C., Seaman, M. S., West, A. P. Jr., Buckley, N., et al. (2015). Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* 522, 487–491. doi: 10.1038/nature14411

Chan, D. C., and Kim, P. S. (1998). HIV entry and its inhibition. *Cell* 93, 681–684. doi: 10.1016/S0092-8674(00)81430-0

Charpentier, C., Nora, T., Tenaillon, O., Clavel, F., and Hance, A. J. (2006). Extensive recombination among human immunodeficiency virus type 1 quasispecies makes an important contribution to viral diversity in individual patients. *J. Virol.* 80, 2472–2482. doi: 10.1128/JVI.80.5.2472-2482.2006

Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P. D., et al. (1996). The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85, 1135–1148. doi: 10.1016/S0092-8674(00)81313-6

Cormier, E. G., and Dragic, T. (2002). The crown and stem of the V3 loop play distinct roles in human immunodeficiency virus type 1 envelope glycoprotein interactions with the CCR5 coreceptor. *J. Virol.* 76, 8953–8957. doi: 10.1128/JVI.76.17.8953-8957.2002

Courter, J. R., Madani, N., Sodroski, J., Schon, A., Freire, E., Kwong, P. D., et al. (2014). Structure-based design, synthesis and validation of CD4-mimetic small molecule inhibitors of HIV-1 entry: conversion of a viral entry agonist to an antagonist. *ACC Chem. Res.* 47, 1228–1237. doi: 10.1021/ar4002735

Curreli, F., Kwon, Y. D., Zhang, H., Yang, Y., Scacalossi, D., Kwong, P. D., et al. (2014). Binding mode characterization of NBD series CD4-mimetic HIV-1 entry inhibitors by X-ray structure and resistance study. *Antimicrob. Agents Chemother.* 58, 5478–5491. doi: 10.1128/AAC.03339-14

Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 312, 763–767. doi: 10.1038/312763a0

Decker, J. M., Bibollet-Ruche, F., Wei, X., Wang, S., Levy, D. N., Wang, W., et al. (2005). Antigenic conservation and immunogenicity of the HIV coreceptor binding site. *J. Exp. Med.* 201, 1407–1419. doi: 10.1084/jem.20042510

Delwart, E. L., Pan, H., Neumann, A., and Markowitz, M. (1998). Rapid, transient changes at the env locus of plasma human immunodeficiency virus type 1 populations during the emergence of protease inhibitor resistance. *J. Virol.* 72, 2416–2421.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., et al. (1996). Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381, 661–666. doi: 10.1038/381661a0

Dereuddre-Bosquet, N., Morellato-Castillo, L., Brouwers, J., Augustijns, P., Bouchemal, K., Ponchel, G., et al. (2012). MiniCD4 microbicide prevents HIV infection of human mucosal explants and vaginal transmission of SHIV(162P3) in cynomolgus macaques. *PLoS Pathog.* 8:e1003071. doi: 10.1371/journal.ppat.1003071

Doranz, B. J., Rucker, J., Yi, Y., Smyth, R. J., Samson, M., Peiper, S. C., et al. (1996). A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* 85, 1149–1158. doi: 10.1016/S0092-8674(00)81314-8

Dorr, P., Westby, M., Dobbs, S., Griffin, P., Irvine, B., Macartney, M., et al. (2005). Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob. Agents Chemother.* 49, 4721–4732. doi: 10.1128/AAC.49.11.4721-4732.2005

Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., et al. (1996). HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 381, 667–673. doi: 10.1038/381667a0

Farzan, M., Mirzabekov, T., Kolchinsky, P., Wyatt, R., Cayabyab, M., Gerard, N. P., et al. (1999). Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell* 96, 667–676. doi: 10.1016/S0092-8674(00)80577-2

Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272, 872–877. doi: 10.1126/science.272.5263.872

Fisher, R. A., Bertonis, J. M., Meier, W., Johnson, V. A., Costopoulos, D. S., Liu, T., et al. (1988). HIV infection is blocked in vitro by recombinant soluble CD4. *Nature* 331, 76–78. doi: 10.1038/331076a0

Fletcher, C. V., Deville, J. G., Samson, P. M., Moye, J. H. Jr., Church, J. A., Spiegel, H. M., et al. (2007). Nonlinear pharmacokinetics of high-dose recombinant fusion protein CD4-IgG2 (PRO 542) observed in HIV-1-infected children. *J. Allergy Clin. Immunol.* 119, 747–750. doi: 10.1016/j.jaci.2006.10.045

Freed, E. O., and Martin, M. A. (1995). The role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection. *J. Biol. Chem.* 270, 23883–23886. doi: 10.1074/jbc.270.41.23883

Ganesh, L., Leung, K., Lore, K., Levin, R., Panet, A., Schwartz, O., et al. (2004). Infection of specific dendritic cells by CCR5-tropic human immunodeficiency virus type 1 promotes cell-mediated transmission of virus resistant to broadly neutralizing antibodies. *J. Virol.* 78, 11980–11987. doi: 10.1128/JVI.78.21.11980-11987.2004

Gardner, M. R., Kattenhorn, L. M., Kondur, H. R., Von Schaewen, M., Dorfman, T., Chiang, J. J., et al. (2015). AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature* 519, 87–91. doi: 10.1038/nature14264

Goulder, P. J., and Watkins, D. I. (2008). Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat. Rev. Immunol.* 8, 619–630. doi: 10.1038/nri2357

Greenberg, M. L., and Cammack, N. (2004). Resistance to enfuvirtide, the first HIV fusion inhibitor. *J. Antimicrob. Chemother.* 54, 333–340. doi: 10.1093/jac/dkh330

Grupping, K., Selhorst, P., Michiels, J., Vereecken, K., Heyndrickx, L., Kessler, P., et al. (2012). MiniCD4 protein resistance mutations affect binding to the HIV-1 gp120 CD4 binding site and decrease entry efficiency. *Retrovirology* 9:36. doi: 10.1186/1742-4690-9-36

Gulick, R. M., Lalezari, J., Goodrich, J., Clumeck, N., Dejesus, E., Horban, A., et al. (2008). Maraviroc for previously treated patients with R5 HIV-1 infection. *N. Engl. J. Med.* 359, 1429–1441. doi: 10.1056/NEJMoa0803152

Haim, H., Si, Z., Madani, N., Wang, L., Courter, J. R., Princiotto, A., et al. (2009). Soluble CD4 and CD4-mimetic compounds inhibit HIV-1 infection by induction of a short-lived activated state. *PLoS Pathog.* 5:e1000360. doi: 10.1371/journal.ppat.1000360

Harada, S., Yoshimura, K., Yamaguchi, A., Boonchawalit, S., Yusa, K., and Matsushita, S. (2013). Impact of antiretroviral pressure on selection of primary

human immunodeficiency virus type 1 envelope sequences in vitro. *J. Gen. Virol.* 94, 933–943. doi: 10.1099/vir.0.047167-0

Huang, C. C., Tang, M., Zhang, M. Y., Majeed, S., Montabana, E., Stanfield, R. L., et al. (2005). Structure of a V3-containing HIV-1 gp120 core. *Science* 310, 1025–1028. doi: 10.1126/science.1118398

Hussey, R. E., Richardson, N. E., Kowalski, M., Brown, N. R., Chang, H. C., Siliciano, R. F., et al. (1988). A soluble CD4 protein selectively inhibits HIV replication and syncytium formation. *Nature* 331, 78–81. doi: 10.1038/331078a0

Ibanez, A., Clotet, B., and Martinez, M. A. (2000). Human immunodeficiency virus type 1 population bottleneck during indinavir therapy causes a genetic drift in the env quasispecies. *J. Gen. Virol.* 81, 85–95. doi: 10.1099/0022-1317-81-1-85

Jacobson, J. M., Lowy, I., Fletcher, C. V., O'neill, T. J., Tran, D. N., Ketas, T. J., et al. (2000). Single-dose safety, pharmacology, and antiviral activity of the human immunodeficiency virus (HIV) type 1 entry inhibitor PRO 542 in HIV-infected adults. *J. Infect. Dis.* 182, 326–329. doi: 10.1086/315698

Kitrinos, K. M., Nelson, J. A., Resch, W., and Swanstrom, R. (2005). Effect of a protease inhibitor-induced genetic bottleneck on human immunodeficiency virus type 1 env gene populations. *J. Virol.* 79, 10627–10637. doi: 10.1128/JVI.79.16.10627-10637.2005

Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., et al. (1984). T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 312, 767–768. doi: 10.1038/312767a0

Kong, R., Xu, K., Zhou, T., Acharya, P., Lemmin, T., Liu, K., et al. (2016). Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* 352, 828–833. doi: 10.1126/science.aae0474

Kuwata, T., Enomoto, I., Baba, M., and Matsushita, S. (2015). Incompatible Natures of the HIV-1 envelope in resistance to the CCR5 antagonist cenicriviroc and to neutralizing antibodies. *Antimicrob. Agents Chemother.* 60, 437–450. doi: 10.1128/AAC.02285-15

Kwon, Y. D., Lalonde, J. M., Yang, Y., Elban, M. A., Sugawara, A., Courter, J. R., et al. (2014). Crystal structures of HIV-1 gp120 envelope glycoprotein in complex with NBD analogues that target the CD4-binding site. *PLoS ONE* 9:e85940. doi: 10.1371/journal.pone.0085940

Kwong, P. D., and Mascola, J. R. (2012). Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. *Immunity* 37, 412–425. doi: 10.1016/j.immuni.2012.08.012

Kwong, P. D., Mascola, J. R., and Nabel, G. J. (2013). Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat. Rev. Immunol.* 13, 693–701. doi: 10.1038/nri3516

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648–659. doi: 10.1038/31405

Lalonde, J. M., Elban, M. A., Courter, J. R., Sugawara, A., Soeta, T., Madani, N., et al. (2011). Design, synthesis and biological evaluation of small molecule inhibitors of CD4-gp120 binding based on virtual screening. *Bioorg. Med. Chem.* 19, 91–101. doi: 10.1016/j.bmc.2010.11.049

Lalonde, J. M., Kwon, Y. D., Jones, D. M., Sun, A. W., Courter, J. R., Soeta, T., et al. (2012). Structure-based design, synthesis, and characterization of dual hotspot small-molecule HIV-1 entry inhibitors. *J. Med. Chem.* 55, 4382–4396. doi: 10.1021/jm300265j

Lalonde, J. M., Le-Khac, M., Jones, D. M., Courter, J. R., Park, J., Schon, A., et al. (2013). Structure-based design and synthesis of an HIV-1 entry inhibitor exploiting X-ray and thermodynamic characterization. *ACS Med. Chem. Lett.* 4, 338–343. doi: 10.1021/ml300407y

Liao, H. X., Lynch, R., Zhou, T., Gao, F., Alam, S. M., Boyd, S. D., et al. (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496, 469–476. doi: 10.1038/nature12053

Lusso, P., Earl, P. L., Sironi, F., Santoro, F., Ripamonti, C., Scarlatti, G., et al. (2005). Cryptic nature of a conserved, CD4-inducible V3 loop neutralization epitope in the native envelope glycoprotein oligomer of CCR5-restricted, but not CXCR4-using, primary human immunodeficiency virus type 1 strains. *J. Virol.* 79, 6957–6968. doi: 10.1128/JVI.79.11.6957-6968.2005

Lynch, R. M., Boritz, E., Coates, E. E., Dezure, A., Madden, P., Costner, P., et al. (2015). Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* 7:319ra206. doi: 10.1126/scitranslmed.aad5752

Madani, N., Perdigoto, A. L., Srinivasan, K., Cox, J. M., Chruma, J. J., Lalonde, J., et al. (2004). Localized changes in the gp120 envelope glycoprotein confer resistance to human immunodeficiency virus entry inhibitors BMS-806 and #155. *J. Virol.* 78, 3742–3752. doi: 10.1128/JVI.78.7.3742-3752.2004

Madani, N., Princiotto, A. M., Easterhoff, D., Bradley, T., Luo, K., Williams, W. B., et al. (2016). Antibodies elicited by multiple envelope glycoprotein immunogens in primates neutralize primary Human Immunodeficiency Viruses (HIV-1) sensitized by CD4-mimetic compounds. *J. Virol.* 90, 5031–5046. doi: 10.1128/JVI.03211-15

Madani, N., Princiotto, A. M., Schon, A., Lalonde, J., Feng, Y., Freire, E., et al. (2014). CD4-mimetic small molecules sensitize human immunodeficiency virus to vaccine-elicited antibodies. *J. Virol.* 88, 6542–6555. doi: 10.1128/JVI.00540-14

Madani, N., Princiotto, A. M., Zhao, C., Jahanbakhshsefidi, F., Mertens, M., Herschhorn, A., et al. (2017). Activation and inactivation of primary human immunodeficiency virus envelope glycoprotein trimers by CD4-mimetic compounds. *J. Virol.* 91:e01880-16. doi: 10.1128/JVI.01880-16

Madani, N., Schon, A., Princiotto, A. M., Lalonde, J. M., Courter, J. R., Soeta, T., et al. (2008). Small-molecule CD4 mimics interact with a highly conserved pocket on HIV-1 gp120. *Structure* 16, 1689–1701. doi: 10.1016/j.str.2008.09.005

Margolis, L., and Shattock, R. (2006). Selective transmission of CCR5-utilizing HIV-1: the 'gatekeeper' problem resolved? *Nat. Rev. Microbiol.* 4, 312–317. doi: 10.1038/nrmicro1427

Matsushita, S., Yoshimura, K., Ramirez, K. P., Pisupati, J., and Murakami, T. (2015). Passive transfer of neutralizing mAb KD-247 reduces plasma viral load in patients chronically infected with HIV-1. *AIDS* 29, 453–462. doi: 10.1097/QAD.0000000000000570

Melillo, B., Liang, S., Park, J., Schon, A., Courter, J. R., Lalonde, J. M., et al. (2016). Small-molecule CD4-mimics: structure-based optimization of HIV-1 entry inhibition. *ACS Med. Chem. Lett.* 7, 330–334. doi: 10.1021/acsmedchemlett.5b00471

Mesplede, T., Moisi, D., Oliveira, M., Ibanescu, I., Ohnona, F., Brenner, B., et al. (2015). Dolutegravir inhibits HIV-1 Env evolution in primary human cells. *AIDS* 29, 659–665. doi: 10.1097/QAD.0000000000000606

Mizuguchi, T., Harada, S., Miura, T., Ohashi, N., Narumi, T., Mori, H., et al. (2016). A minimally cytotoxic CD4 mimic as an HIV entry inhibitor. *Bioorg. Med. Chem. Lett.* 26, 397–400. doi: 10.1016/j.bmcl.2015.11.103

Moore, P. L., Gray, E. S., Wibmer, C. K., Bhiman, J. N., Nonyane, M., Sheward, D. J., et al. (2012). Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* 18, 1688–1692. doi: 10.1038/nm.2985

Myszka, D. G., Sweet, R. W., Hensley, P., Brigham-Burke, M., Kwong, P. D., Hendrickson, W. A., et al. (2000). Energetics of the HIV gp120-CD4 binding reaction. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9026–9031. doi: 10.1073/pnas.97.16.9026

Narumi, T., Arai, H., Yoshimura, K., Harada, S., Hirota, Y., Ohashi, N., et al. (2013). CD4 mimics as HIV entry inhibitors: lead optimization studies of the aromatic substituents. *Bioorg. Med. Chem.* 21, 2518–2526. doi: 10.1016/j.bmc.2013.02.041

Narumi, T., Arai, H., Yoshimura, K., Harada, S., Nomura, W., Matsushita, S., et al. (2011). Small molecular CD4 mimics as HIV entry inhibitors. *Bioorg. Med. Chem.* 19, 6735–6742. doi: 10.1016/j.bmc.2011.09.045

Narumi, T., Ochiai, C., Yoshimura, K., Harada, S., Tanaka, T., Nomura, W., et al. (2010). CD4 mimics targeting the HIV entry mechanism and their hybrid molecules with a CXCR4 antagonist. *Bioorg. Med. Chem. Lett.* 20, 5853–5858. doi: 10.1016/j.bmcl.2010.07.106

Nijhuis, M., Boucher, C. A., Schipper, P., Leitner, T., Schuurman, R., and Albert, J. (1998). Stochastic processes strongly influence HIV-1 evolution during suboptimal protease-inhibitor therapy. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14441–14446. doi: 10.1073/pnas.95.24.14441

Nora, T., Charpentier, C., Tenaillon, O., Hoede, C., Clavel, F., and Hance, A. J. (2007). Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment. *J. Virol.* 81, 7620–7628. doi: 10.1128/JVI.00083-07

Ohashi, N., Harada, S., Mizuguchi, T., Irahara, Y., Yamada, Y., Kotani, M., et al. (2016). Small-molecule CD4 mimics containing mono-cyclohexyl moieties as HIV entry inhibitors. *ChemMedChem* 11, 940–946. doi: 10.1002/cmdc.201500590

Pfaff, J. M., Wilen, C. B., Harrison, J. E., Demarest, J. F., Lee, B., Doms, R. W., et al. (2010). HIV-1 resistance to CCR5 antagonists associated with highly efficient use of CCR5 and altered tropism on primary CD4+ T cells. *J. Virol.* 84, 6505–6514. doi: 10.1128/JVI.00374-10

Pugach, P., Ketas, T. J., Michael, E., and Moore, J. P. (2008). Neutralizing antibody and anti-retroviral drug sensitivities of HIV-1 isolates resistant to small molecule CCR5 inhibitors. *Virology* 377, 401–407. doi: 10.1016/j.virol.2008.04.032

Pugach, P., Marozsan, A. J., Ketas, T. J., Landes, E. L., Moore, J. P., and Kuhmann, S. E. (2007). HIV-1 clones resistant to a small molecule CCR5 inhibitor use the inhibitor-bound form of CCR5 for entry. *Virology* 361, 212–228. doi: 10.1016/j.virol.2006.11.004

Pugach, P., Ray, N., Klasse, P. J., Ketas, T. J., Michael, E., Doms, R. W., et al. (2009). Inefficient entry of vicriviroc-resistant HIV-1 via the inhibitor-CCR5 complex at low cell surface CCR5 densities. *Virology* 387, 296–302. doi: 10.1016/j.virol.2009.02.044

Putcharoen, O., Lee, S. H., Henrich, T. J., Hu, Z., Vanichanan, J., Coakley, E., et al. (2012). HIV-1 clinical isolates resistant to CCR5 antagonists exhibit delayed entry kinetics that are corrected in the presence of drug. *J. Virol.* 86, 1119–1128. doi: 10.1128/JVI.06421-11

Rambaut, A., Posada, D., Crandall, K. A., and Holmes, E. C. (2004). The causes and consequences of HIV evolution. *Nat. Rev. Genet.* 5, 52–61. doi: 10.1038/nrg1246

Rangel, H. R., Weber, J., Chakraborty, B., Gutierrez, A., Marotta, M. L., Mirza, M., et al. (2003). Role of the human immunodeficiency virus type 1 envelope gene in viral fitness. *J. Virol.* 77, 9069–9073. doi: 10.1128/JVI.77.16.9069-9073.2003

Ratcliff, A. N., Shi, W., and Arts, E. J. (2013). HIV-1 resistance to maraviroc conferred by a CD4 binding site mutation in the envelope glycoprotein gp120. *J. Virol.* 87, 923–934. doi: 10.1128/JVI.01863-12

Raymond, S., Maillard, A., Amiel, C., Peytavin, G., Trabaud, M. A., Desbois, D., et al. (2015). Virological failure of patients on maraviroc-based antiretroviral therapy. *J. Antimicrob. Chemother.* 70, 1858–1864. doi: 10.1093/jac/dkv026

Reeves, J. D., Gallo, S. A., Ahmad, N., Miamidian, J. L., Harvey, P. E., Sharron, M., et al. (2002). Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics. *Proc. Natl. Acad. Sci. U.S.A.* 99, 16249–16254. doi: 10.1073/pnas.252469399

Richard, J., Veillette, M., Brassard, N., Iyer, S. S., Roger, M., Martin, L., et al. (2015). CD4 mimetics sensitize HIV-1-infected cells to ADCC. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2687–E2694. doi: 10.1073/pnas.1506755112

Robertson, D. (2003). US FDA approves new class of HIV therapeutics. *Nat. Biotechnol.* 21, 470–471. doi: 10.1038/nbt0503-470

Roche, M., Jakobsen, M. R., Ellett, A., Salimiseyedabad, H., Jubb, B., Westby, M., et al. (2011). HIV-1 predisposed to acquiring resistance to maraviroc (MVC) and other CCR5 antagonists in vitro has an inherent, low-level ability to utilize MVC-bound CCR5 for entry. *Retrovirology* 8:89. doi: 10.1186/1742-4690-8-89

Roche, M., Salimi, H., Duncan, R., Wilkinson, B. L., Chikere, K., Moore, M. S., et al. (2013). A common mechanism of clinical HIV-1 resistance to the CCR5 antagonist maraviroc despite divergent resistance levels and lack of common gp120 resistance mutations. *Retrovirology* 10:43. doi: 10.1186/1742-4690-10-43

Sattentau, Q. J., and Moore, J. P. (1995). Human immunodeficiency virus type 1 neutralization is determined by epitope exposure on the gp120 oligomer. *J. Exp. Med.* 182, 185–196. doi: 10.1084/jem.182.1.185

Schon, A., Madani, N., Klein, J. C., Hubicki, A., Ng, D., Yang, X., et al. (2006). Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry* 45, 10973–10980. doi: 10.1021/bi061193r

Seaman, M. S., Janes, H., Hawkins, N., Grandpre, L. E., Devoy, C., Giri, A., et al. (2010). Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol.* 84, 1439–1452. doi: 10.1128/JVI.02108-09

Sheehy, N., Desselberger, U., Whitwell, H., and Ball, J. K. (1996). Concurrent evolution of regions of the envelope and polymerase genes of human immunodeficiency virus type 1 during zidovudine (AZT) therapy. *J. Gen. Virol.* 77(Pt 5), 1071–1081. doi: 10.1099/0022-1317-77-5-1071

Smith, D. H., Byrn, R. A., Marsters, S. A., Gregory, T., Groopman, J. E., and Capon, D. J. (1987). Blocking of HIV-1 infectivity by a soluble, secreted form of the CD4 antigen. *Science* 238, 1704–1707. doi: 10.1126/science.3500514

Starcich, B. R., Hahn, B. H., Shaw, G. M., Mcneely, P. D., Modrow, S., Wolf, H., et al. (1986). Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 45, 637–648. doi: 10.1016/0092-8674(86)90778-6

Thali, M., Moore, J. P., Furman, C., Charles, M., Ho, D. D., Robinson, J., et al. (1993). Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* 67, 3978–3988.

Tilton, J. C., Wilen, C. B., Didigu, C. A., Sinha, R., Harrison, J. E., Agrawal-Gamse, C., et al. (2010). A maraviroc-resistant HIV-1 with narrow cross-resistance to other CCR5 antagonists depends on both N-terminal and extracellular loop domains of drug-bound CCR5. *J. Virol.* 84, 10863–10876. doi: 10.1128/JVI.01109-10

van Gils, M. J., Van Den Kerkhof, T. L., Ozorowski, G., Cottrell, C. A., Sok, D., Pauthner, M., et al. (2016). An HIV-1 antibody from an elite neutralizer implicates the fusion peptide as a site of vulnerability. *Nat. Microbiol.* 2:16199. doi: 10.1038/nmicrobiol.2016.199

Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., et al. (1999). Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13091–13096. doi: 10.1073/pnas.96.23.13091

Westby, M., Lewis, M., Whitcomb, J., Youle, M., Pozniak, A. L., James, I. T., et al. (2006). Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. *J. Virol.* 80, 4909–4920. doi: 10.1128/JVI.80.10.4909-4920.2006

Westby, M., Smith-Burchnell, C., Mori, J., Lewis, M., Mosley, M., Stockdale, M., et al. (2007). Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *J. Virol.* 81, 2359–2371. doi: 10.1128/JVI.02006-06

Wyatt, R., and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280, 1884–1888. doi: 10.1126/science.280.5371.1884

Yamada, Y., Ochiai, C., Yoshimura, K., Tanaka, T., Ohashi, N., Narumi, T., et al. (2010). CD4 mimics targeting the mechanism of HIV entry. *Bioorg. Med. Chem. Lett.* 20, 354–358. doi: 10.1016/j.bmcl.2009.10.098

Yoshimura, K., Harada, S., Boonchawalit, S., Kawanami, Y., and Matsushita, S. (2014). Impact of maraviroc-resistant and low-CCR5-adapted mutations induced by in vitro passage on sensitivity to anti-envelope neutralizing antibodies. *J. Gen. Virol.* 95, 1816–1826. doi: 10.1099/vir.0.062885-0

Yoshimura, K., Harada, S., Shibata, J., Hatada, M., Yamada, Y., Ochiai, C., et al. (2010). Enhanced exposure of human immunodeficiency virus type 1 primary isolate neutralization epitopes through binding of CD4 mimetic compounds. *J. Virol.* 84, 7558–7568. doi: 10.1128/JVI.00227-10

Yoshimura, K., Shibata, J., Kimura, T., Honda, A., Maeda, Y., Koito, A., et al. (2006). Resistance profile of a neutralizing anti-HIV monoclonal antibody, KD-247, that shows favourable synergism with anti-CCR5 inhibitors. *AIDS* 20, 2065–2073. doi: 10.1097/01.aids.0000247587.31320.fe

Zhang, Y. M., Dawson, S. C., Landsman, D., Lane, H. C., and Salzman, N. P. (1994). Persistence of four related human immunodeficiency virus subtypes during the course of zidovudine therapy: relationship between virion RNA and proviral DNA. *J. Virol.* 68, 425–432.

Zhao, Q., Ma, L., Jiang, S., Lu, H., Liu, S., He, Y., et al. (2005). Identification of N-phenyl-N'-(2,2,6,6-tetramethyl-piperidin-4-yl)-oxalamides as a new class of HIV-1 entry inhibitors that prevent gp120 binding to CD4. *Virology* 339, 213–225. doi: 10.1016/j.virol.2005.06.008

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Harada and Yoshimura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

GENERAL COMMENTARY
published: 24 February 2016
doi: 10.3389/fmicb.2016.00254

Commentary: MARCH8 Inhibits HIV-1 Infection by Reducing Virion Incorporation of Envelope Glycoproteins

Mikako Fujita *

Research Institute for Drug Discovery, School of Pharmacy, Kumamoto University, Kumamoto, Japan

Keywords: MARCH8, SAMHD1, macrophage, HIV, reservoir

A commentary on

MARCH8 inhibits HIV-1 infection by reducing virion incorporation of envelope glycoproteins
by Tada, T., Zhang, Y., Koyama, T., Tobiume, M., Tsunetsugu-Yokota, Y., Yamaoka, S., et al. (2015). Nat. Med. 21, 1502–1507. doi: 10.1038/nm.3956

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University Graduate School, Japan

Reviewed by:
Ai Kawana-Tachikawa,
University of Tokyo, Japan
Takamasa Ueno,
Kumamoto University, Japan
Shinya Suzu,
Kumamoto University, Japan

***Correspondence:**
Mikako Fujita
mfujita@kumamoto-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *23 January 2016*
Accepted: *15 February 2016*
Published: *24 February 2016*

Citation:
Fujita M (2016) Commentary: MARCH8 Inhibits HIV-1 Infection by Reducing Virion Incorporation of Envelope Glycoproteins. Front. Microbiol. 7:254. doi: 10.3389/fmicb.2016.00254

Recently, Kenzo Tokunaga's group reported a novel restriction factor against HIV, MARCH8, which is highly expressed in terminally differentiated myeloid cells such as macrophages. Virus infection in macrophages was first observed in HIV-infected patients in the mid-1980s (Gyorkey et al., 1985; Ho et al., 1986; Koenig et al., 1986). Three decades have passed since then; however, the role of HIV-infected macrophages in AIDS pathogenesis remains controversial. Here, some potential implications of Tokunaga et al.'s study on this controversy will be addressed.

SIV-infected rhesus monkeys are a good model for investigating the role of HIV-infected macrophages because their pathology resembles the slow progression of AIDS in humans. A comparative study of rhesus macaques infected with T cell-tropic SIVmac239 (Kestler et al., 1990) or with macrophage-tropic SIVmac316 (Mori et al., 1992), which carries nine mutations compared with SIVmac239 (Johnson et al., 2003) found that SIVmac316 replicates with the simian body as well as SIVmac239 just after inoculation. However, SIVmac316 induces a slower disease progression than SIVmac239, demonstrating that the contribution of virus-infected macrophages to pathogenesis is smaller than that of virus-infected T cells.

Studies have also used an SIV that lacks expression of its accessory protein, Vpx, which is critical for SIV/HIV-2 replication in macrophages and resting T lymphocytes and is also important in activated T lymphocytes (Fujita et al., 2010, 2012; Baldauf et al., 2012). Rhesus macaques infected with a *vpx*-deleted SIVmac239 eventually died after a slower disease progression than that of animals infected with wild-type SIVmac239 (Westmoreland et al., 2014). In monkeys infected with *vpx*-deleted SIVmac239, minimal macrophage infection was detected, even though infected macrophages were observed following wild-type SIV infection.

Furthermore, there was a recent study of rhesus macaques infected with SIVmac239 or SIVmac316 mutants, both of which had mutations in Vpx inhibiting the ability of this protein to confer infectivity. The viruses that recovered their replication ability in this study only appeared in the animals infected with the T cell-tropic SIVmac239, demonstrating the lower importance of virus replication in macrophages than that in T cells (Shingai et al., 2015). Based on these results, it is likely that HIV has difficulty replicating in macrophages *in vivo* and that HIV-infected macrophages play a minimal role in progressing the general symptom of AIDS.

Although HIV-infected macrophages are not critical for disease progression, their role in HIV infection may be to serve as an HIV reservoir in the body, acting as an obstacle to HIV

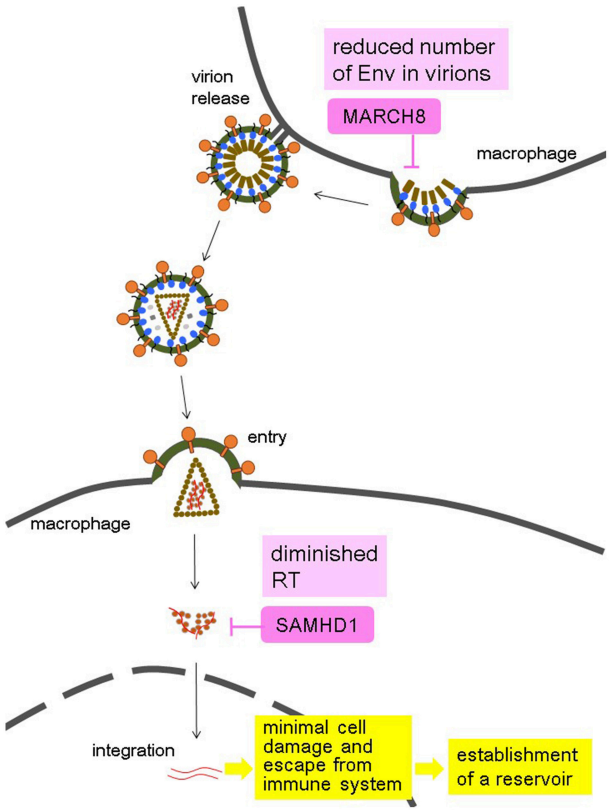


FIGURE 1 | Hypothesis of the HIV-1 strategy to use macrophages as a reservoir by taking advantage of the synergistic effects of MARCH8 and SAMHD1 in macrophages. MARCH8 reduces the virion incorporation of Env proteins into producer cells, and SAMHD1 diminishes the amount of reverse transcription (RT) in target cells, both of which lead to minimal cell damage induced by a mild amount of HIV-1 replication, and escape from immune system. These allow macrophages to become an HIV reservoir.

eradication by antiretroviral therapy (ART). The existence of an HIV reservoir has been postulated since just after the establishment of ART (Chun and Fauci, 1999), and although a 2001 report suggested that it might be composed of macrophages (Igarashi et al., 2001), until recently (Churchill et al., 2016) it was generally believed to be composed of memory $CD4^+$ T cells, partially because it is difficult to elucidate which cell type(s) constitute the reservoir by using patients or through laboratory experiments. In support of the hypothesis that macrophages are the HIV reservoir, HIV-infected macrophages were observed in HIV-infected patients with undetectable plasma viral loads (Cribbs et al., 2015). Even taking into account that macrophages are resistant to HIV replication, macrophages may serve as part of long-lived HIV reservoir.

Great progress into understanding HIV-resistance in macrophages, specifically the discovery of two host restriction factors in macrophages, has recently been made. One of these restriction factors is SAMHD1 (Hrecka et al., 2011; Laguette et al., 2011). This protein was found as a target protein of Vpx, and it reduces reverse transcription (RT) products. Investigations into the function of SAMHD1 initially focused on its dNTPase activity (Goldstone et al., 2011; Powell et al., 2011), reducing dNTP pools, materials of genomic cDNA (Kim et al., 2012; Lahouassa et al., 2012). However, it was later proposed that SAMHD1 uses its RNase activity to degrade HIV RNA before reverse transcription (Beloglazova et al., 2013; Ryoo et al., 2014). It is presently unclear if one of these or both are responsible for the activity of SAMHD1 (Ballana and Esté, 2015). Interestingly, the HIV-2 Vpx protein is able to degrade SAMHD1 (Hrecka et al., 2011; Laguette et al., 2011), while HIV-1 lacks a special protein to combat SAMHD1. Although the reverse transcriptase of HIV-1 is more efficient than that of HIV-2 (Lenzi et al., 2015), the artificial incorporation of Vpx into HIV-1 virions dramatically increases their infectivity in macrophages (Goujon et al., 2008), showing that HIV-1 does not sufficiently overcome the function of SAMHD1.

Another recently discovered host factor in macrophages is membrane-associated RING-CH8 (MARCH8) (Tada et al., 2015). This protein has been known to downregulate various transmembrane proteins. As with many great scientific discoveries, the identification of MARCH8 as a macrophage host factor began with a serendipitous finding. Tada et al. initially noticed that MARCH8-expressing lentiviral vectors had a low infectivity and later found that a large amount of MARCH8 is specifically expressed in terminally differentiated myeloid cells, macrophages, and dendritic cells. MARCH8 was demonstrated to drastically reduce HIV-1 virion incorporation of envelope glycoproteins and inhibit its infectivity. The same inhibitory effect was observed in virions containing envelope proteins from HIV-2, SIV, MLV, or VSV. MARCH8 was suggested to interact with HIV-1 Env, leading to its downregulation from surface of producer cells. Interestingly, neither HIV-1 Vpr, Vpu nor Nef have detectable anti-MARCH8 activity, suggesting that HIV-1 lacks a mechanism to directly combat the effects of MARCH8.

HIV, particularly HIV-1, may have evolved a way of taking advantage of the effects of host restriction proteins such as SAMHD1 and MARCH8 (**Figure 1**). The synergistic suppression of infectivity by these factors and other effects likely leads to a mild amount of HIV replication in macrophages, causing minimal cellular damage. Furthermore, virus could escape from host immune system. These permit virus survival. The long life of these cells allows them to serve as viral reservoirs, present even in patients with undetectable plasma viral loads after receiving ART. Future studies should aim to devise ways of targeting the macrophage reservoir cells to fully eliminate HIV.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

Baldauf, H. M., Pan, X., Erikson, E., Schmidt, S., Daddacha, W., Burggraf, M., et al. (2012). SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nat. Med.* 18, 1682–1687. doi: 10.1038/nm.2964

Ballana, E., and Esté, J. A. (2015). SAMHD1: at the crossroads of cell proliferation, immune responses, and virus restriction. *Trends Microbiol.* 23, 680–692. doi: 10.1016/j.tim.2015.08.002

Beloglazova, N., Flick, R., Tchigvintsev, A., Brown, G., Popovic, A., Nocek, B., et al. (2013). Nuclease activity of the human SAMHD1 protein implicated in the Aicardi-Goutieres syndrome and HIV-1 restriction. *J. Biol. Chem.* 288, 8101–8110. doi: 10.1074/jbc.M112.431148

Chun, T. W., and Fauci, A. S. (1999). Latent reservoirs of HIV: obstacles to the eradication of virus. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10958–10961. doi: 10.1073/pnas.96.20.10958

Churchill, M. J., Deeks, S. G., Margolis, D. M., Siliciano, R. F., and Swanstrom, R. (2016). HIV reservoirs: what, where and how to target them. *Nat. Rev. Microbiol.* 14, 55–60. doi: 10.1038/nrmicro.2015.5

Cribbs, S. K., Lennox, J., Caliendo, A. M., Brown, L. A., and Guidot, D. M. (2015). Healthy HIV-1-infected individuals on highly active antiretroviral therapy harbor HIV-1 in their alveolar macrophages. *AIDS Res. Hum. Retroviruses* 31, 64–70. doi: 10.1089/aid.2014.0133

Fujita, M., Nomaguchi, M., Adachi, A., and Otsuka, M. (2012). SAMHD1-Dependent and -independent functions of HIV-2/SIV Vpx protein. *Front. Microbiol.* 3:297. doi: 10.3389/fmicb.2012.00297

Fujita, M., Otsuka, M., Nomaguchi, M., and Adachi, A. (2010). Multifaceted activity of HIV Vpr/Vpx proteins: the current view of their virological functions. *Rev. Med. Virol.* 20, 68–76. doi: 10.1002/rmv.636

Goldstone, D. C., Ennis-Adeniran, V., Hedden, J. J., Groom, H. C., Rice, G. I., Christodoulou, E., et al. (2011). HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* 480, 379–382. doi: 10.1038/nature10623

Goujon, C., Arfi, V., Pertel, T., Luban, J., Lienard, J., Rigal, D., et al. (2008). Characterization of simian immunodeficiency virus SIVSM/human immunodeficiency virus type 2 Vpx function in human myeloid cells. *J. Virol.* 82, 12335–12345. doi: 10.1128/JVI.01181-08

Gyorkey, F., Melnick, J. L., Sinkovics, J. G., and Gyorkey, P. (1985). Retrovirus resembling HTLV in macrophages of patients with AIDS. *Lancet* 1, 106. doi: 10.1016/S0140-6736(85)91995-6

Ho, D. D., Rota, T. R., and Hirsch, M. S. (1986). Infection of monocyte/macrophages by human T lymphotropic virus type III. *J. Clin. Invest.* 77, 1712–1715. doi: 10.1172/JCI112491

Hrecka, K., Hao, C., Gierszewska, M., Swanson, S. K., Kesik-Brodacka, M., Srivastava, S., et al. (2011). Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474, 658–661. doi: 10.1038/nature10195

Igarashi, T., Brown, C. R., Endo, Y., Buckler-White, A., Plishka, R., Bischofberger, N., et al. (2001). Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): Implications for HIV-1 infections of humans. *Proc. Natl. Acad. Sci. U.S.A.* 98, 658–663. doi: 10.1073/pnas.98.2.658

Johnson, W. E., Lifson, J. D., Lang, S. M., Johnson, R. P., and Desrosiers, R. C. (2003). Importance of B-cell responses for immunological control of variant strains of simian immunodeficiency virus. *J. Virol.* 77, 375–381. doi: 10.1128/JVI.77.1.375-381.2003

Kestler, H., Kodama, T., Ringler, D., Marthas, M., Pedersen, N., Lackner, A., et al. (1990). Induction of AIDS in rhesus monkeys by molecularly cloned simian immunodeficiency virus. *Science* 248, 1109–1112. doi: 10.1126/science.2160735

Kim, B., Nguyen, L. A., Daddacha, W., and Hollenbaugh, J. A. (2012). Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *J. Biol. Chem.* 287, 21570–21574. doi: 10.1074/jbc.C112.374843

Koenig, S., Gendelman, H. E., Orenstein, J. M., Dal Canto, M. C., Pezeshkpour, G. H., Yungbluth, M., et al. (1986). Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy. *Science* 233, 1089–1093. doi: 10.1126/science.3016903

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Ségéral, E., et al. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474, 654–657. doi: 10.1038/nature10117

Lahouassa, H., Daddacha, W., Hofmann, H., Ayinde, D., Logue, E. C., Dragin, L., et al. (2012). SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* 13, 223–228. doi: 10.1038/ni.2236

Lenzi, G. M., Domaoal, R. A., Kim, D. H., Schinazi, R. F., and Kim, B. (2015). Mechanistic and kinetic differences between reverse transcriptases of Vpx coding and non-coding lentiviruses. *J. Biol. Chem.* 290, 30078–30086. doi: 10.1074/jbc.M115.691576

Mori, K., Ringler, D. J., Kodama, T., and Desrosiers, R. C. (1992). Complex determinants of macrophage tropism in env of simian immunodeficiency virus. *J. Virol.* 66, 2067–2075.

Powell, R. D., Holland, P. J., Hollis, T., and Perrino, F. W. (2011). Aicardi-Goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. *J. Biol. Chem.* 286, 43596–43600. doi: 10.1074/jbc.C111.317628

Ryoo, J., Choi, J., Oh, C., Kim, S., Seo, M., Kim, S. Y., et al. (2014). The ribonuclease activity of SAMHD1 is required for HIV-1 restriction. *Nat. Med.* 20, 936–941. doi: 10.1038/nm.3626

Shingai, M., Welbourn, S., Brenchley, J. M., Acharya, P., Miyagi, E., Plishka, R. J., et al. (2015). The expression of functional Vpx during pathogenic SIVmac infections of rhesus macaques suppresses SAMHD1 in CD4+ memory T cells. *PLoS Pathog.* 11:e1004928. doi: 10.1371/journal.ppat.1004928

Tada, T., Zhang, Y., Koyama, T., Tobiume, M., Tsunetsugu-Yokota, Y., Yamaoka, S., et al. (2015). MARCH8 inhibits HIV-1 infection by reducing virion incorporation of envelope glycoproteins. *Nat. Med.* 21, 1502–1507. doi: 10.1038/nm.3956

Westmoreland, S. V., Converse, A. P., Hrecka, K., Hurley, M., Knight, H., Piatak, M., et al. (2014). SIV vpx is essential for macrophage infection but not for development of AIDS. *PLoS ONE* 9:e84463. doi: 10.1371/journal.pone.0084463

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer TU declared a shared affiliation, though no other collaboration, with the author MF to the handling Editor, who ensured that the process nevertheless met the standards of a fair and objective review. The reviewer SS declared a shared affiliation, though no other collaboration, with the author MF to the handling Editor, who ensured that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Fujita. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

MINI REVIEW
published: 19 December 2016
doi: 10.3389/fmicb.2016.02027

APOBEC3G-Mediated G-to-A Hypermutation of the HIV-1 Genome: The Missing Link in Antiviral Molecular Mechanisms

*Ayaka Okada[1] and Yasumasa Iwatani[1,2]**

[1] Department of Microbiology and Immunology, Laboratory of Infectious Diseases, Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan, [2] Department of AIDS Research, Nagoya University Graduate School of Medicine, Nagoya, Japan

APOBEC3G (A3G) is a member of the cellular polynucleotide cytidine deaminases, which catalyze the deamination of cytosine (dC) to uracil (dU) in single-stranded DNA. These enzymes potently inhibit the replication of a variety of retroviruses and retrotransposons, including HIV-1. A3G is incorporated into *vif*-deficient HIV-1 virions and targets viral reverse transcripts, particularly minus-stranded DNA products, in newly infected cells. It is well established that the enzymatic activity of A3G is closely correlated with the potential to greatly inhibit HIV-1 replication in the absence of Vif. However, the details of the underlying molecular mechanisms are not fully understood. One potential mechanism of A3G antiviral activity is that the A3G-dependent deamination may trigger degradation of the dU-containing reverse transcripts by cellular uracil DNA glycosylases (UDGs). More recently, another mechanism has been suggested, in which the virion-incorporated A3G generates lethal levels of the G-to-A hypermutation in the viral DNA genome, thus potentially driving the viruses into "error catastrophe" mode. In this mini review article, we summarize the deaminase-dependent and deaminase-independent molecular mechanisms of A3G and discuss how A3G-mediated deamination is linked to antiviral mechanisms.

Keywords: HIV-1, APOBEC3G, antiviral mechanisms, deamination, reverse transcription

OPEN ACCESS

Edited by:
Yasuko Tsunetsugu Yokota, Tokyo University of Technology, Japan

Reviewed by:
Hiroaki Takeuchi, Tokyo Medical and Dental University, Japan
Jean-Christophe Paillart, Centre National de la Recherche Scientifique – University of Strasbourg, France

****Correspondence:***
Yasumasa Iwatani
iwataniy@nnh.hosp.go.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *03 November 2016*
Accepted: *02 December 2016*
Published: *19 December 2016*

Citation:
Okada A and Iwatani Y (2016) APOBEC3G-Mediated G-to-A Hypermutation of the HIV-1 Genome: The Missing Link in Antiviral Molecular Mechanisms. Front. Microbiol. 7:2027. doi: 10.3389/fmicb.2016.02027

INTRODUCTION

The human apolipoprotein B mRNA editing enzyme catalytic subunit 3 (A3) family of cellular polynucleotide cytidine deaminases comprises seven members (A, B, C, D, F, G, and H) that catalyze the conversion of cytosine (dC) to uracil (dU) on single-stranded DNA (ssDNA; Kitamura et al., 2011). These enzymes, particularly A3G, exhibit potent antiviral activity against retrotransposons and retroviruses, including HIV-1 (Sheehy et al., 2002; Simon et al., 2005; Bogerd et al., 2006; Muckenfuss et al., 2006; Okeoma et al., 2007; Hultquist et al., 2011; Chaipan et al., 2013). However, the HIV-1 accessory protein viral infectivity factor (Vif), antagonizes the A3G-mediated host defense system, thereby promoting the propagation of HIV-1 in human cells (Sheehy et al., 2002; Conticello et al., 2003). Vif recruits A3G into the E3 ubiquitin ligase complex, including Cullin5, ElonginB/C, and core binding factor subunit beta (CBFβ), and promotes A3G degradation

through the ubiquitin-proteasome pathway (Yu et al., 2003; Yu Y. et al., 2004; Salter et al., 2012; Guo et al., 2014). Thus, in the absence of Vif, A3G is incorporated into HIV-1 virions from their viral producers and exerts its antiviral activity in newly infected cells (Sheehy et al., 2002; Anderson and Hope, 2008). The A3G incorporation depends on its binding affinity to the viral nucleocapsid (NC) domain of Gag and/or to the viral/non-viral RNAs (Cen et al., 2004; Luo et al., 2004; Svarovskaia et al., 2004; Burnett and Spearman, 2007; Strebel and Khan, 2008; Apolonia et al., 2015; York et al., 2016). In infected cells, A3G inhibits viral replication through the specific deamination of dCs in viral minus-strand DNA, thus resulting in massive G-to-A hypermutation of the nascent viral DNA (vDNA) genome during reverse transcription (Lecossier et al., 2003; Zhang et al., 2003; Suspene et al., 2004). The A3G-induced hypermutation is observed as a discrete "all or nothing" phenomenon (Armitage et al., 2012). In addition, A3G directly blocks reverse transcriptase (RT) elongation in a deaminase-independent manner (Guo et al., 2006, 2007; Iwatani et al., 2007; Bishop et al., 2008; Gillick et al., 2013; Chaurasiya et al., 2014) and interferes with the integration of proviral DNA into the host chromosome (Luo et al., 2007; Mbisa et al., 2007, 2010). These cooperative molecular mechanisms are likely to be important in maximizing the anti-HIV-1 activity of A3G. Nevertheless, several studies have shown that the enzymatic activity of A3G is closely correlated with the potential to highly inhibit *vif*-deficient HIV-1 replication (Mangeat et al., 2003; Navarro et al., 2005; Iwatani et al., 2006; Browne et al., 2009). In contrast to understanding of deaminase-independent mechanisms, the details of deaminase-dependent mechanisms, in which A3G inhibits *vif*-deficient HIV-1 replication, are not fully understood.

Unique Features of A3G-Mediated Deamination

The N-terminal and C-terminal domains (NTD and CTD, respectively) of A3G both contain Zn coordinate motifs ((H/C)xE(x)$_{23-28}$PCxxC; Wedekind et al., 2003; Conticello et al., 2005). The A3G CTD is catalytically active, whereas its NTD has no enzymatic activity but exhibits strong binding to ssDNA and RNA (Hache et al., 2005; Navarro et al., 2005; Iwatani et al., 2006). During the reverse transcription of *vif*-deficient HIV-1, A3G preferentially deaminates the second dC of 5′-CC dinucleotide sites in the newly synthesized viral minus-stranded ssDNA (Harris et al., 2003; Mangeat et al., 2003; Zhang et al., 2003; Yu Q. et al., 2004). This dinucleotide preference is unique among A3 family proteins (Hultquist et al., 2011; Rathore et al., 2013). This deamination occurs more efficiently at the dC close to the 5′-end of ssDNA and less efficiently at the last ∼30 nt of the 3′ ssDNA end, the so-called dead zone (Chelico et al., 2006, 2008). Therefore, it is likely that A3G more efficiently catalyzes the deamination of ssDNA when the A3G CTD is oriented toward the 5′ ssDNA end, and the A3G NTD restricts access of the CTD to the dead zone (Chelico et al., 2010; Shlyakhtenko et al., 2015). Furthermore, the deamination efficacy decreases with decreasing ssDNA length (Chelico et al., 2006), thus probably reflecting the infrequent orientation of the A3G CTD toward the 5′ ssDNA end (Shlyakhtenko et al., 2015).

Deaminase-Dependent Antiviral Mechanisms

Error Catastrophe

APOBEC3G deaminase activity is crucial for its antiviral activity and restriction of *vif*-deficient HIV-1 replication (Mangeat et al., 2003; Navarro et al., 2005; Iwatani et al., 2006; Browne et al., 2009). An experimental-mathematical study estimated that 99.3% of the antiviral effect of A3G is dependent on its deaminase activity (Kobayashi et al., 2014) (**Figure 1**). Many reports have consistently supported the presumable deaminase-dependent mechanism in which massive A3G-mediated hypermutations in viral reverse transcripts cause lethal mutational loads that terminate progeny virus production and subsequent virus propagation (Harris et al., 2003; Lecossier et al., 2003; Mangeat et al., 2003; Zhang et al., 2003; Suspene et al., 2004; Rawson et al., 2015). This mechanism has previously been described as the error catastrophe mechanism (Crotty et al., 2001; Eigen, 2002; Graci and Cameron, 2002). The mutations introduced in the viral genome, to a certain threshold, lead to sequence diversification, thus enabling adaptation to environmental changes. In contrast, massive amounts of mutations caused by mutagens lead to viral replication failure, called error catastrophe. A3G excessively converts dC to dU in the vDNA of *vif*-deficient HIV-1, thus resulting in G-to-A hypermutations in the viral integrated genomes. These mutations include substitutions of tryptophan codons to in-frame premature stop codons and/or may introduce amino acid changes lethal for viral replication. A3G probably hinders functional viral protein expression and progeny virus production (Pace et al., 2006) (**Figure 1**). A recent study has demonstrated that the introduction of C-to-U mutations in the trans-activation response (TAR) element, a key regulation factor of HIV-1 transcription elongation, results in an early block of viral gene expression (Nowarski et al., 2014) (**Figure 1**).

APOBEC3G-induced G-to-A hypermutations that block viral replication have also been detected in the proviral genomes of infected patients, probably because natural variants of Vif do not completely neutralize A3G and other A3 family proteins (Janini et al., 2001; Kieffer et al., 2005; Simon et al., 2005). Individuals with low to undetectable plasma HIV-1 RNA levels, referred to as "elite controllers," have frequent G-to-A hypermutations (Gandhi et al., 2008). Therefore, A3G is also presumably involved in the production of defective viruses *in vivo* (Pace et al., 2006; Armitage et al., 2012; Eyzaguirre et al., 2013; Krisko et al., 2013; Sato et al., 2014; Delviks-Frankenberry et al., 2016). However, the incomplete neutralization of A3 family proteins by Vif might result in sequence diversification *in vivo* (Simon et al., 2005; Kim et al., 2010, 2014; Sato et al., 2014; Alteri et al., 2015).

Degradation of Uracilated DNA

APOBEC3G, compared with catalytically inactive A3G, decreases the copy number of reverse transcripts in the early phases of infection (Anderson and Hope, 2008; Bishop et al., 2008). In addition, before A3G was identified as a Vif-related cellular factor, von Schwedler et al. (1993) had reported that levels

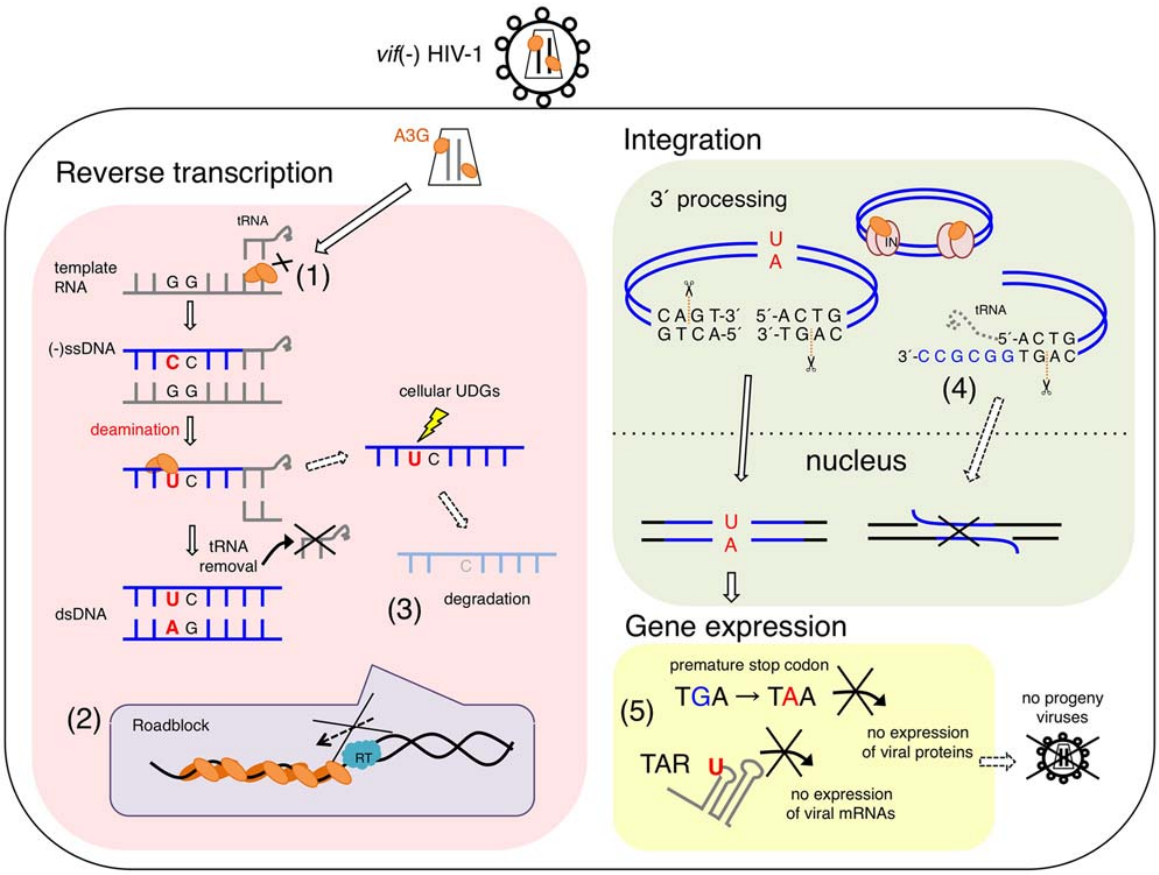


FIGURE 1 | Proposed antiviral mechanisms of APOBEC3G (A3G) in *vif*-deficient HIV-1-infected cells. A3G potentially blocks several processes of *vif*-deficient HIV-1 [*vif*(–) HIV-1] replication in a infected cell. (1) A3G interferes with the annealing of the tRNA primer for reverse transcription. (2) A3G physically blocks reverse transcription elongation by binding to RNA and/or ssDNA templates, referred to as a "roadblock" mechanism. A3G oligomerization is closely associated with the efficiency of this action. (3) A3G-induced C-to-U mutations trigger the degradation of reverse transcripts by cellular uracil DNA glycosylases (UNG1, etc.). (3) Direct interaction of A3G with HIV-1 integrase blocks the formation of the replication-competent pre-integration DNA complex. (4) A3G decreases the efficiency and specificity of tRNA primer removal, thereby producing a poor substrate for strand transfer and integration. (5) dC-to-dU conversions in viral (–)ssDNA cause G-to-A hypermutations in progeny viral genomes, thereby leading to viral replication failure, called "error catastrophe." In part, dC-to-dU mutations in the trans-activation response (TAR) element result in an early block of HIV-1 transcription.

of the reverse transcripts of *vif*-deficient HIV-1 are decreased in newly infected cells when the virus is produced from non-permissive cell lines (currently known as cell lines expressing high amounts of A3G). Thus, it was initially proposed that A3G-induced C-to-U mutations in nascent reverse transcripts might trigger the degradation of reverse transcripts by cellular uracil DNA glycosylases (UDGs), such as nuclear UNG2 and SMUG1 (Harris et al., 2003) (**Figure 1**). The UDG-mediated removal of uracil bases from reverse transcripts might result in the digestion of DNA products at the abasic site by apurinic/apyrimidinic endonuclease. One study further supporting this possibility has shown that the antiviral activity of A3G is partially affected by the UNG2 inhibitor (Ugi) and siRNA specific to UNG2 in virus-producing cells but not in target cells (Yang et al., 2007). However, other studies have shown that UNG2 and SMUG1 are dispensable for the antiviral activity of A3G: A3G-mediated antiviral activity is not changed by Ugi expression (Kaiser and Emerman, 2006; Mbisa et al., 2007; Langlois and Neuberger, 2008), and A3G activity has been observed in Epstein-Barr virus-transformed B-cell lines derived from a *UNG2*$^{-/-}$ patient (Kaiser and Emerman, 2006) and in a SMUG1-deficient avian cell line, with or without exogenous Ugi expression (Langlois and Neuberger, 2008). More recently, two studies have shown involvement of uracilated vDNAs in their chromosomal integration during infection of human cells that contain high levels of dUTP. Yan et al. (2011) have reported that the uracilated vDNA protected it from autointegration, which resulted in facilitating chromosomal integration and viral replication. In contrast, the other study by Hansen et al. (2016) indicated that heavily uracilated vDNAs in monocyte-derived macrophages, not in T-lymphocytes, were not efficiently integrated into chromosomal DNA due to their UNG2-dependent degradation in the nucleus. These data suggest different fate of uracilated vDNA between cytoplasm and nucleus during HIV-1 infection. In addition, because the deaminase-dependent antiviral mechanism has been observed

in a variety of cell types, unidentified cellular factors might determine the fate of vDNA containing A3G-induced uracil. Therefore, additional studies are required to determine whether other cellular uracil DNA repair enzymes beyond UNG2 and SMUG1, are involved in the degradation of nascent reverse transcripts.

Deaminase-Independent Antiviral Mechanisms

Although A3G-mediated deamination was initially proposed to be the sole mechanism of the antiviral activity against *vif*-deficient HIV-1, subsequent studies have demonstrated that other mechanisms are also involved in the inhibition of viral replication. In addition, the enzymatic activity of A3F is not absolutely required for its inhibitory effect on *vif*-deficient HIV-1 replication (Holmes et al., 2007; Luo et al., 2007; Mbisa et al., 2010). Furthermore, a deaminase activity-deficient A3G mutant blocks the replication of HIV-1, mouse mammary tumor virus, and murine leukemia virus, to a certain extent (Okeoma et al., 2007; Belanger et al., 2013), thus suggesting the broad specificity of antiviral activity in terms of the deaminase-independent mechanism.

Initially, Guo et al. (2006, 2007) suggested that A3G might interfere with $tRNA^{Lys3}$ primer placement in viral reverse transcription, in a manner independent of A3G-mediated deamination (**Figure 1**). However, such inhibition of primer annealing has not been observed in other studies (Iwatani et al., 2007; Bishop et al., 2008). Instead, the inhibition of HIV-1 RT elongation has been demonstrated by using *in vitro* and *in vivo* systems (Iwatani et al., 2007; Bishop et al., 2008; Adolph et al., 2013; Belanger et al., 2013) (**Figure 1**). It has been suggested that the inhibitory effect reflects the following unique biochemical characteristics of A3G: (1) A3G protein exhibits high affinity binding specifically to single-stranded polynucleotides, such as ssDNA and RNA (Iwatani et al., 2006; Polevoda et al., 2015); (2) A3G, compared with RT, exhibits significantly higher binding affinity for polynucleotides, although A3G shows similar or slightly less binding affinity for ssDNA than the NC (Iwatani et al., 2006; Darlix et al., 2011); (3) A3G mediates homo-oligomerization in a dose-dependent manner in the presence of ssDNA or RNA, whereas A3G forms monomers, dimers, and tetramers in the absence of these polynucleotides (Wedekind et al., 2006; Salter et al., 2009); and (4) A3G initially binds ssDNA with rapid on-off rates and subsequently converts to a slow dissociation mode after homo-oligomerization (Chaurasiya et al., 2014). Therefore, A3G probably inhibits reverse transcription by tightly binding to the ssDNA or RNA template, thus forming a roadblock that physically obstructs viral DNA synthesis (Iwatani et al., 2007; Adolph et al., 2013; Chaurasiya et al., 2014) (**Figure 1**). This deaminase-independent mechanism might increase the availability of ssDNA for deamination by A3G (Adolph et al., 2013; Chaurasiya et al., 2014), thereby resulting in cooperative effects between deaminase-dependent and deaminase-independent mechanisms.

A3G-mediated inhibition of plus-strand DNA transfer and integration has also been observed (Mbisa et al., 2007, 2010) (**Figure 1**). A3G decreases the efficiency and specificity of tRNA processing and removal during reverse transcription, thereby producing aberrant viral DNA ends defective for efficient plus-strand transfer and integration. Interestingly, it has been reported that A3F exerts an inhibitory effect on viral DNA integration, although its mechanism differs from that of A3G; A3F prevents integration by its binding to the double-stranded DNA of the proviral DNA ends (Mbisa et al., 2010). In contrast to the competition of nucleic acid interactions between A3G and RT/integrase, direct interactions of the A3G protein with HIV-1 RT (Wang et al., 2012) or integrase (Luo et al., 2007) (**Figure 1**) have been reported to be a deaminase-independent mechanism, although the molecular mechanism underlying the specific affinity of A3G for a variety of retroviral RTs and integrases remains unclear. This may be associated with a loss of the reverse transcription complex structure in newly infected cells when A3G coexists with RT (Carr et al., 2006).

Structural Basis of Antiviral Mechanisms

Recent progress in determining the A3G protein structure has enhanced the current understanding of A3G-mediated antiviral mechanisms, particularly interactions between nucleotides and A3G. First, three-dimensional structures of the A3G CTD were determined by using NMR spectroscopy (Chen et al., 2008; Furukawa et al., 2009; Harjes et al., 2009) and X-ray crystallography (Holden et al., 2008; Shandilya et al., 2010; Li et al., 2012; Lu et al., 2015). Although the structure of the A3G CTD/ssDNA complex has not yet been determined experimentally, four different structural models of the ssDNA-bound A3G CTD have been proposed to explain how A3G recognizes the ssDNA substrate. The first ssDNA-bound model shows the nearly vertical orientation of ssDNA relative to helices α2 and α3 of the A3G CTD along a cleft around the Zn-coordination center (Zn-center pocket; "brim" model) (Chen et al., 2008). The second model suggests that ssDNA binds to the Zn-center pocket with the ssDNA crossed over the cleft seen in the brim model ("kinked" model; Holden et al., 2008). The third model resembles the brim model, although in this model, helices α2 and α3 are involved in ssDNA binding to a greater extent ("straight" model; Furukawa et al., 2009). The recently proposed fourth model based on the crystal structure of the A3G CTD-CTD dimer is a hybrid model of the kinked and brim models (Lu et al., 2015). Nevertheless, it remains inconclusive which ssDNA substrate-binding model is appropriate for deamination catalysis.

Structures of the highly insoluble A3G NTD protein have recently been determined by using NMR spectroscopy (Kouno et al., 2015) and X-ray crystallography (Xiao et al., 2016). The crystal structure of the A3G NTD, derived from the rhesus macaque (*Macaca mulatta*) protein, reveals a detailed structural mechanism illustrating A3G dimerization and the interaction between the A3G NTD and ssDNA (Xiao et al., 2016). The structural data suggest that ssDNA binding to the A3G NTD changes the conformation of the loops around the Zn-center pocket and Y124 in loop7, thus functioning as a "molecular switch" that regulates the opened/closed status of the Zn-center pocket. The structure also indicates that the dimerization interfaces of the A3G NTD dimer provide a large positively charged surface, including the Zn-center pocket,

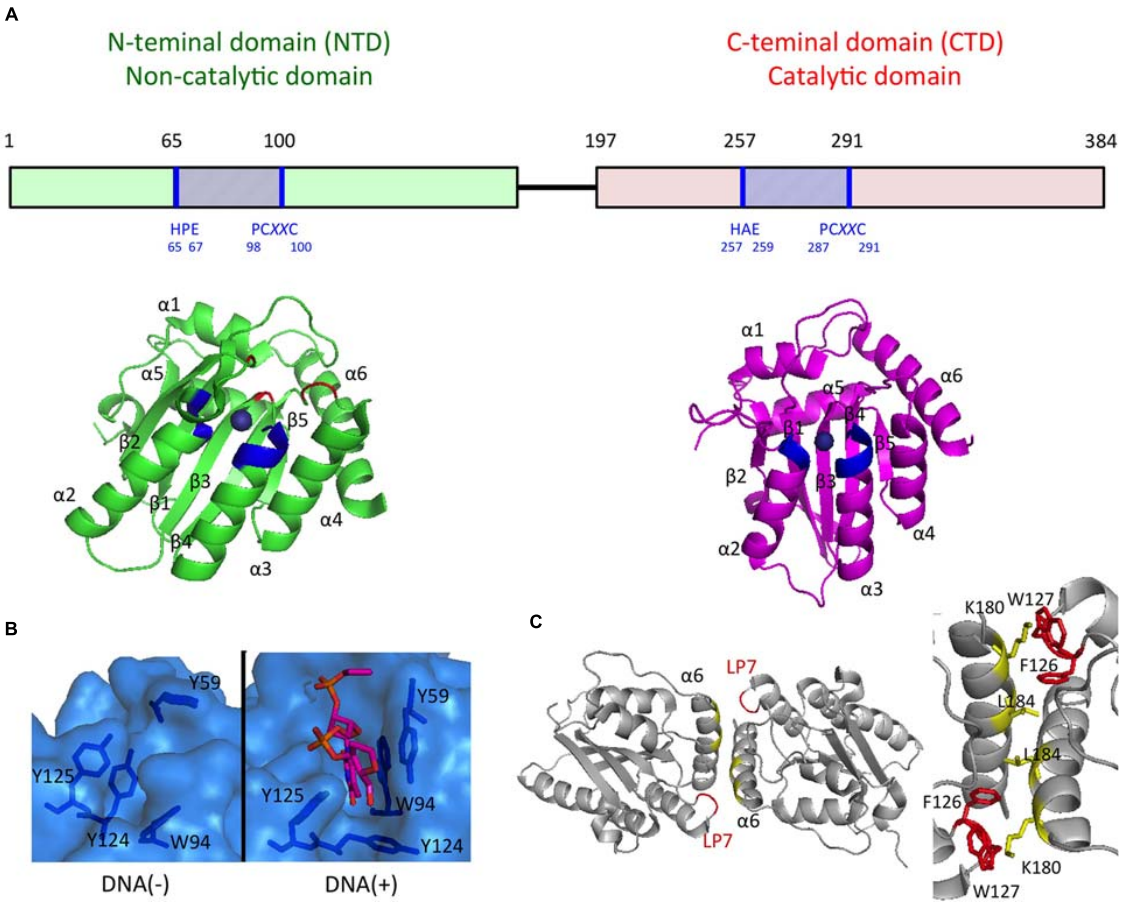


FIGURE 2 | APOBEC3G domain structures important for the antiviral activities. (A) Schematics of the A3G domain primary sequence and the domain structures are shown in the Upper and Lower, respectively. Zn-coordination motifs and the main DNA/RNA binding sites surrounding Zn are blue and red, respectively. **(B)** The surface representation of the DNA/RNA binding site in the absence (Left) or presence (Right) of a ssDNA substrate. **(C)** Structure of the rhesus A3G-NTD dimer (Left) and key residues involved in dimerization interfaces (Right).

thereby resulting in the formation of a high affinity surface toward the ssDNA or RNA (**Figure 2**). These structural features are consistent with results of previous biochemical studies suggesting that the NTD-NTD interaction is crucial for A3G oligomerization, nucleic acid binding, and the antiviral activity of A3G (Bennett et al., 2008; Huthoff et al., 2009; Chelico et al., 2010; Shandilya et al., 2010; Belanger et al., 2013; Chaurasiya et al., 2014).

CONCLUSION

Recent evidence suggests that A3G executes potent antiviral activity through cooperative deaminase-dependent and deaminase-independent mechanisms. Undoubtedly, the enzymatic activity of A3G is closely correlated with the potential to inhibit *vif*-deficient HIV-1 replication. However, it remains unclear how the A3G-mediated deamination event is linked to the A3G-mediated lethal inhibition of viral replication. Further studies of the molecular mechanisms of A3G antiviral activity, particularly for the deaminase-dependent mechanisms, are required, including the careful determination of the fate of uracil-containing viral DNA in newly HIV-1-infected cells.

AUTHOR CONTRIBUTIONS

AO and YI analyzed the data and wrote the paper.

FUNDING

This work was financially supported in part by the Japan Society for the Promotion of Science KAKENHI [grant number 15H04740 (to YI)].

ACKNOWLEDGMENT

We thank Dr. Hirotaka Ode (National Hospital Organization, Nagoya Medical Center) for helpful discussions.

REFERENCES

Adolph, M. B., Webb, J., and Chelico, L. (2013). Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase. *PLoS ONE* 8:e64196. doi: 10.1371/journal.pone.0064196

Alteri, C., Surdo, M., Bellocchi, M. C., Saccomandi, P., Continenza, F., Armenia, D., et al. (2015). Incomplete APOBEC3G/F neutralization by HIV-1 Vif mutants facilitates the genetic evolution from CCR5 to CXCR4 usage. *Antimicrob. Agents Chemother.* 59, 4870–4881. doi: 10.1128/AAC.00137-15

Anderson, J. L., and Hope, T. J. (2008). APOBEC3G restricts early HIV-1 replication in the cytoplasm of target cells. *Virology* 375, 1–12. doi: 10.1016/j.virol.2008.01.042

Apolonia, L., Schulz, R., Curk, T., Rocha, P., Swanson, C. M., Schaller, T., et al. (2015). Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS Pathog.* 11:e1004609. doi: 10.1371/journal.ppat.1004609

Armitage, A. E., Deforche, K., Chang, C. H., Wee, E., Kramer, B., Welch, J. J., et al. (2012). APOBEC3G-induced hypermutation of human immunodeficiency virus type-1 is typically a discrete "all or nothing" phenomenon. *PLoS Genet* 8:e1002550. doi: 10.1371/journal.pgen.1002550

Belanger, K., Savoie, M., Rosales Gerpe, M. C., Couture, J. F., and Langlois, M. A. (2013). Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic Acids Res.* 41, 7438–7452. doi: 10.1093/nar/gkt527

Bennett, R. P., Salter, J. D., Liu, X., Wedekind, J. E., and Smith, H. C. (2008). APOBEC3G subunits self-associate via the C-terminal deaminase domain. *J. Biol. Chem.* 283, 33329–33336. doi: 10.1074/jbc.M803726200

Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M., and Malim, M. H. (2008). APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.* 4:e1000231. doi: 10.1371/journal.ppat.1000231

Bogerd, H. P., Wiegand, H. L., Hulme, A. E., Garcia-Perez, J. L., O'Shea, K. S., Moran, J. V., et al. (2006). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8780–8785. doi: 10.1073/pnas.0603313103

Browne, E. P., Allers, C., and Landau, N. R. (2009). Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. *Virology* 387, 313–321. doi: 10.1016/j.virol.2009.02.026

Burnett, A., and Spearman, P. (2007). APOBEC3G multimers are recruited to the plasma membrane for packaging into human immunodeficiency virus type 1 virus-like particles in an RNA-dependent process requiring the NC basic linker. *J. Virol.* 81, 5000–5013. doi: 10.1128/JVI.02237-06

Carr, J. M., Davis, A. J., Coolen, C., Cheney, K., Burrell, C. J., and Li, P. (2006). Vif-deficient HIV reverse transcription complexes (RTCs) are subject to structural changes and mutation of RTC-associated reverse transcription products. *Virology* 351, 80–91. doi: 10.1016/j.virol.2006.03.027

Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004). The interaction between HIV-1 Gag and APOBEC3G. *J. Biol. Chem.* 279, 33177–33184. doi: 10.1074/jbc.M402062200

Chaipan, C., Smith, J. L., Hu, W. S., and Pathak, V. K. (2013). APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *J. Virol.* 87, 444–453. doi: 10.1128/JVI.00676-12

Chaurasiya, K. R., McCauley, M. J., Wang, W., Qualley, D. F., Wu, T., Kitamura, S., et al. (2014). Oligomerization transforms human APOBEC3G from an efficient enzyme to a slowly dissociating nucleic acid-binding protein. *Nat. Chem.* 6, 28–33. doi: 10.1038/nchem.1795

Chelico, L., Pham, P., Calabrese, P., and Goodman, M. F. (2006). APOBEC3G DNA deaminase acts processively 3' – > 5' on single-stranded DNA. *Nat. Struct. Mol. Biol.* 13, 392–399. doi: 10.1038/nsmb1086

Chelico, L., Prochnow, C., Erie, D. A., Chen, X. S., and Goodman, M. F. (2010). Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G. *J. Biol. Chem.* 285, 16195–16205. doi: 10.1074/jbc.M110.107987

Chelico, L., Sacho, E. J., Erie, D. A., and Goodman, M. F. (2008). A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. *J. Biol. Chem.* 283, 13780–13791. doi: 10.1074/jbc.M801004200

Chen, K. M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., et al. (2008). Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* 452, 116–119. doi: 10.1038/nature06638

Conticello, S. G., Harris, R. S., and Neuberger, M. S. (2003). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr. Biol.* 13, 2009–2013. doi: 10.1016/j.cub.2003.10.034

Conticello, S. G., Thomas, C. J., Petersen-Mahrt, S. K., and Neuberger, M. S. (2005). Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol.* 22, 367–377. doi: 10.1093/molbev/msi026

Crotty, S., Cameron, C. E., and Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6895–6900. doi: 10.1073/pnas.111085598

Darlix, J. L., Godet, J., Ivanyi-Nagy, R., Fosse, P., Mauffret, O., and Mely, Y. (2011). Flexible nature and specific functions of the HIV-1 nucleocapsid protein. *J. Mol. Biol.* 410, 565–581. doi: 10.1016/j.jmb.2011.03.037

Delviks-Frankenberry, K. A., Nikolaitchik, O. A., Burdick, R. C., Gorelick, R. J., Keele, B. F., Hu, W. S., et al. (2016). Minimal contribution of APOBEC3-induced G-to-A hypermutation to HIV-1 recombination and genetic variation. *PLoS Pathog* 12:e1005646. doi: 10.1371/journal.ppat.1005646

Eigen, M. (2002). Error catastrophe and antiviral strategy. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13374–13376. doi: 10.1073/pnas.212514799

Eyzaguirre, L. M., Charurat, M., Redfield, R. R., Blattner, W. A., Carr, J. K., and Sajadi, M. M. (2013). Elevated hypermutation levels in HIV-1 natural viral suppressors. *Virology* 443, 306–312. doi: 10.1016/j.virol.2013.05.019

Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., et al. (2009). Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* 28, 440–451. doi: 10.1038/emboj.2008.290

Gandhi, S. K., Siliciano, J. D., Bailey, J. R., Siliciano, R. F., and Blankson, J. N. (2008). Role of APOBEC3G/F-mediated hypermutation in the control of human immunodeficiency virus type 1 in elite suppressors. *J. Virol.* 82, 3125–3130. doi: 10.1128/JVI.01533-07

Gillick, K., Pollpeter, D., Phalora, P., Kim, E. Y., Wolinsky, S. M., and Malim, M. H. (2013). Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *J. Virol.* 87, 1508–1517. doi: 10.1128/JVI.02587-12

Graci, J. D., and Cameron, C. E. (2002). Quasispecies, error catastrophe, and the antiviral activity of ribavirin. *Virology* 298, 175–180. doi: 10.1006/viro.2002.1487

Guo, F., Cen, S., Niu, M., Saadatmand, J., and Kleiman, L. (2006). Inhibition of tRNA(3)(Lys)-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *J. Virol.* 80, 11710–11722. doi: 10.1128/JVI.01038-06

Guo, F., Cen, S., Niu, M., Yang, Y., Gorelick, R. J., and Kleiman, L. (2007). The interaction of APOBEC3G with human immunodeficiency virus type 1 nucleocapsid inhibits tRNA3Lys annealing to viral RNA. *J. Virol.* 81, 11322–11331. doi: 10.1128/JVI.00162-07

Guo, Y., Dong, L., Qiu, X., Wang, Y., Zhang, B., Liu, H., et al. (2014). Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* 505, 229–233. doi: 10.1038/nature12884

Hache, G., Liddament, M. T., and Harris, R. S. (2005). The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J. Biol. Chem.* 280, 10920–10924. doi: 10.1074/jbc.M500382200

Hansen, E. C., Ransom, M., Hesselberth, J. R., Hosmane, N. N., Capoferri, A. A., Bruner, K. M., et al. (2016). Diverse fates of uracilated HIV-1 DNA during infection of myeloid lineage cells. *Elife* 5:e18447. doi: 10.7554/eLife.18447

Harjes, E., Gross, P. J., Chen, K. M., Lu, Y., Shindo, K., Nowarski, R., et al. (2009). An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J. Mol. Biol.* 389, 819–832. doi: 10.1016/j.jmb.2009.04.031

Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., et al. (2003). DNA deamination mediates innate immunity to retroviral infection. *Cell* 113, 803–809. doi: 10.1016/S0092-8674(03)00423-9

Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., et al. (2008). Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* 456, 121–124. doi: 10.1038/nature07357

Holmes, R. K., Koning, F. A., Bishop, K. N., and Malim, M. H. (2007). APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.* 282, 2587–2595. doi: 10.1074/jbc.M607298200

Hultquist, J. F., Lengyel, J. A., Refsland, E. W., LaRue, R. S., Lackey, L., Brown, W. L., et al. (2011). Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J. Virol.* 85, 11220–11234. doi: 10.1128/JVI.05238-11

Huthoff, H., Autore, F., Gallois-Montbrun, S., Fraternali, F., and Malim, M. H. (2009). RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathog.* 5:e1000330. doi: 10.1371/journal.ppat.1000330

Iwatani, Y., Chan, D. S., Wang, F., Maynard, K. S., Sugiura, W., Gronenborn, A. M., et al. (2007). Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.* 35, 7096–7108. doi: 10.1093/nar/gkm750

Iwatani, Y., Takeuchi, H., Strebel, K., and Levin, J. G. (2006). Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.* 80, 5992–6002. doi: 10.1128/JVI.02680-05

Janini, M., Rogers, M., Birx, D. R., and McCutchan, F. E. (2001). Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J. Virol.* 75, 7973–7986. doi: 10.1128/JVI.75.17.7973-7986.2001

Kaiser, S. M., and Emerman, M. (2006). Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G. *J. Virol.* 80, 875–882. doi: 10.1128/JVI.80.2.875-882.2006

Kieffer, T. L., Kwon, P., Nettles, R. E., Han, Y., Ray, S. C., and Siliciano, R. F. (2005). G– > A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J. Virol.* 79, 1975–1980. doi: 10.1128/JVI.79.3.1975-1980.2005

Kim, E. Y., Bhattacharya, T., Kunstman, K., Swantek, P., Koning, F. A., Malim, M. H., et al. (2010). Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. *J. Virol.* 84, 10402–10405. doi: 10.1128/JVI.01223-10

Kim, E. Y., Lorenzo-Redondo, R., Little, S. J., Chung, Y. S., Phalora, P. K., Maljkovic Berry, I., et al. (2014). Human APOBEC3 induced mutation of human immunodeficiency virus type-1 contributes to adaptation and evolution in natural infection. *PLoS Pathog.* 10:e1004281. doi: 10.1371/journal.ppat.1004281

Kitamura, S., Ode, H., and Iwatani, Y. (2011). Structural features of antiviral APOBEC3 proteins are linked to their functional activities. *Front Microbiol* 2:258. doi: 10.3389/fmicb.2011.00258

Kobayashi, T., Koizumi, Y., Takeuchi, J. S., Misawa, N., Kimura, Y., Morita, S., et al. (2014). Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *J. Virol.* 88, 5881–5887. doi: 10.1128/JVI.00062-14

Kouno, T., Luengas, E. M., Shigematsu, M., Shandilya, S. M., Zhang, J., Chen, L., et al. (2015). Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. *Nat. Struct. Mol. Biol.* 22, 485–491. doi: 10.1038/nsmb.3033

Krisko, J. F., Martinez-Torres, F., Foster, J. L., and Garcia, J. V. (2013). HIV restriction by APOBEC3 in humanized mice. *PLoS Pathog.* 9:e1003242. doi: 10.1371/journal.ppat.1003242

Langlois, M. A., and Neuberger, M. S. (2008). Human APOBEC3G can restrict retroviral infection in avian cells and acts independently of both UNG and SMUG1. *J. Virol.* 82, 4660–4664. doi: 10.1128/JVI.02469-07

Lecossier, D., Bouchonnet, F., Clavel, F., and Hance, A. J. (2003). Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* 300: 1112. doi: 10.1126/science.1083338

Li, M., Shandilya, S. M., Carpenter, M. A., Rathore, A., Brown, W. L., Perkins, A. L., et al. (2012). First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. *ACS Chem. Biol.* 7, 506–517. doi: 10.1021/cb200440y

Lu, X., Zhang, T., Xu, Z., Liu, S., Zhao, B., Lan, W., et al. (2015). Crystal structure of DNA cytidine deaminase ABOBEC3G catalytic deamination domain suggests a binding mode of full-length enzyme to single-stranded DNA. *J. Biol. Chem.* 290, 4010–4021. doi: 10.1074/jbc.M114.624262

Luo, K., Liu, B., Xiao, Z., Yu, Y., Yu, X., Gorelick, R., et al. (2004). Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. *J. Virol.* 78, 11841–11852. doi: 10.1128/JVI.78.21.11841-11852.2004

Luo, K., Wang, T., Liu, B., Tian, C., Xiao, Z., Kappes, J., et al. (2007). Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. *J. Virol.* 81, 7238–7248. doi: 10.1128/JVI.02584-06

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424, 99–103. doi: 10.1038/nature01709

Mbisa, J. L., Barr, R., Thomas, J. A., Vandegraaff, N., Dorweiler, I. J., Svarovskaia, E. S., et al. (2007). Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J. Virol.* 81, 7099–7110. doi: 10.1128/JVI.00272-07

Mbisa, J. L., Bu, W., and Pathak, V. K. (2010). APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J. Virol.* 84, 5250–5259. doi: 10.1128/JVI.02358-09

Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Lower, J., Cichutek, K., et al. (2006). APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.* 281, 22161–22172. doi: 10.1074/jbc.M601716200

Navarro, F., Bollman, B., Chen, H., Konig, R., Yu, Q., Chiles, K., et al. (2005). Complementary function of the two catalytic domains of APOBEC3G. *Virology* 333, 374–386. doi: 10.1016/j.virol.2005.01.011

Nowarski, R., Prabhu, P., Kenig, E., Smith, Y., Britan-Rosich, E., and Kotler, M. (2014). APOBEC3G inhibits HIV-1 RNA elongation by inactivating the viral trans-activation response element. *J. Mol. Biol.* 426, 2840–2853. doi: 10.1016/j.jmb.2014.05.012

Okeoma, C. M., Lovsin, N., Peterlin, B. M., and Ross, S. R. (2007). APOBEC3 inhibits mouse mammary tumour virus replication in vivo. *Nature* 445, 927–930. doi: 10.1038/nature05540

Pace, C., Keller, J., Nolan, D., James, I., Gaudieri, S., Moore, C., et al. (2006). Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J. Virol.* 80, 9259–9269. doi: 10.1128/JVI.00888-06

Polevoda, B., McDougall, W. M., Tun, B. N., Cheung, M., Salter, J. D., Friedman, A. E., et al. (2015). RNA binding to APOBEC3G induces the disassembly of functional deaminase complexes by displacing single-stranded DNA substrates. *Nucleic Acids Res.* 43, 9434–9445. doi: 10.1093/nar/gkv970

Rathore, A., Carpenter, M. A., Demir, O., Ikeda, T., Li, M., Shaban, N. M., et al. (2013). The local dinucleotide preference of APOBEC3G can be altered from 5′-CC to 5′-TC by a single amino acid substitution. *J. Mol. Biol.* 425, 4442–4454. doi: 10.1016/j.jmb.2013.07.040

Rawson, J. M., Landman, S. R., Reilly, C. S., and Mansky, L. M. (2015). HIV-1 and HIV-2 exhibit similar mutation frequencies and spectra in the absence of G-to-A hypermutation. *Retrovirology* 12: 60. doi: 10.1186/s12977-015-0180-6

Salter, J. D., Krucinska, J., Raina, J., Smith, H. C., and Wedekind, J. E. (2009). A hydrodynamic analysis of APOBEC3G reveals a monomer-dimer-tetramer self-association that has implications for anti-HIV function. *Biochemistry* 48, 10685–10687. doi: 10.1021/bi901642c

Salter, J. D., Lippa, G. M., Belashov, I. A., and Wedekind, J. E. (2012). Core-binding factor beta increases the affinity between human Cullin 5 and HIV-1 Vif within an E3 ligase complex. *Biochemistry* 51, 8702–8704. doi: 10.1021/bi301244z

Sato, K., Takeuchi, J. S., Misawa, N., Izumi, T., Kobayashi, T., Kimura, Y., et al. (2014). APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. *PLoS Pathog.* 10:e1004453. doi: 10.1371/journal.ppat.1004453

Shandilya, S. M., Nalam, M. N., Nalivaika, E. A., Gross, P. J., Valesano, J. C., Shindo, K., et al. (2010). Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure* 18, 28–38. doi: 10.1016/j.str.2009.10.016

Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646–650. doi: 10.1038/nature00939

Shlyakhtenko, L. S., Dutta, S., Banga, J., Li, M., Harris, R. S., and Lyubchenko, Y. L. (2015). APOBEC3G Interacts with ssDNA by Two Modes: AFM Studies. *Sci. Rep.* 5, 15648. doi: 10.1038/srep15648

Simon, V., Zennou, V., Murray, D., Huang, Y., Ho, D. D., and Bieniasz, P. D. (2005). Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog.* 1:e6. doi: 10.1371/journal.ppat.0010006

Strebel, K., and Khan, M. A. (2008). APOBEC3G encapsidation into HIV-1 virions: which RNA is it? *Retrovirology* 5:55. doi: 10.1186/1742-4690-5-55

Suspene, R., Sommer, P., Henry, M., Ferris, S., Guetard, D., Pochet, S., et al. (2004). APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase. *Nucleic Acids Res.* 32, 2421–2429. doi: 10.1093/nar/gkh554

Svarovskaia, E. S., Xu, H., Mbisa, J. L., Barr, R., Gorelick, R. J., Ono, A., et al. (2004). Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J. Biol. Chem.* 279, 35822–35828. doi: 10.1074/jbc.M405761200

von Schwedler, U., Song, J., Aiken, C., and Trono, D. (1993). Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *J. Virol.* 67, 4945–4955.

Wang, X., Ao, Z., Chen, L., Kobinger, G., Peng, J., and Yao, X. (2012). The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. *J. Virol.* 86, 3777–3786. doi: 10.1128/JVI.06594-11

Wedekind, J. E., Dance, G. S., Sowden, M. P., and Smith, H. C. (2003). Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet.* 19, 207–216. doi: 10.1016/S0168-9525(03)00054-4

Wedekind, J. E., Gillilan, R., Janda, A., Krucinska, J., Salter, J. D., Bennett, R. P., et al. (2006). Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. *J. Biol. Chem.* 281, 38122–38126. doi: 10.1074/jbc.C600253200

Xiao, X., Li, S. X., Yang, H., and Chen, X. S. (2016). Crystal structures of APOBEC3G N-domain alone and its complex with DNA. *Nat. Commun.* 7:12193. doi: 10.1038/ncomms12193

Yan, N., O'Day, E., Wheeler, L. A., Engelman, A., and Lieberman, J. (2011). HIV DNA is heavily uracilated, which protects it from autointegration. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9244–9249. doi: 10.1073/pnas.1102943108

Yang, B., Chen, K., Zhang, C., Huang, S., and Zhang, H. (2007). Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J. Biol. Chem.* 282, 11667–11675. doi: 10.1074/jbc.M606864200

York, A., Kutluay, S. B., Errando, M., and Bieniasz, P. D. (2016). The RNA binding specificity of human APOBEC3 proteins resembles that of HIV-1 nucleocapsid. *PLoS Pathog.* 12:e1005833. doi: 10.1371/journal.ppat.1005833

Yu, Q., Konig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., et al. (2004). Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* 11, 435–442. doi: 10.1038/nsmb758

Yu, Y., Xiao, Z., Ehrlich, E. S., Yu, X., and Yu, X. F. (2004). Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev.* 18, 2867–2872. doi: 10.1101/gad.1250204

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., et al. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302, 1056–1060. doi: 10.1126/science.1089591

Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C., and Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424, 94–98. doi: 10.1038/nature01707

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Okada and Iwatani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 August 2016
doi: 10.3389/fmicb.2016.01211

Expression Profiles of Vpx/Vpr Proteins Are Co-related with the Primate Lentiviral Lineage

Yosuke Sakai[1], Ariko Miyake[2], Naoya Doi[1], Hikari Sasada[1], Yasuyuki Miyazaki[3], Akio Adachi[1] and Masako Nomaguchi[1]**

[1] Department of Microbiology, Tokushima University Graduate School of Medical Science, Tokushima, Japan, [2] Laboratory of Molecular Immunology and Infectious Disease, Joint Faculty of Veterinary Medicine, Yamaguchi University, Yamaguchi, Japan, [3] Department of Microbiology and Cell Biology, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

Viruses of human immunodeficiency virus type 2 (HIV-2) and some simian immunodeficiency virus (SIV) lineages carry a unique accessory protein called Vpx. Vpx is essential or critical for viral replication in natural target cells such as macrophages and T lymphocytes. We have previously shown that a poly-proline motif (PPM) located at the C-terminal region of Vpx is required for its efficient expression in two strains of HIV-2 and SIVmac, and that the Vpx expression levels of the two clones are significantly different. Notably, the PPM sequence is conserved and confined to Vpx and Vpr proteins derived from certain lineages of HIV-2/SIVs. In this study, Vpx/Vpr proteins from diverse primate lentiviral lineages were experimentally and phylogenetically analyzed to obtain the general expression picture in cells. While both the level and PPM-dependency of Vpx/Vpr expression in transfected cells varied among viral strains, each viral group, based on Vpx/Vpr amino acid sequences, was found to exhibit a characteristic expression profile. Moreover, phylogenetic tree analyses on Gag and Vpx/Vpr proteins gave essentially the same results. Taken together, our study described here suggests that each primate lentiviral lineage may have developed a unique expression pattern of Vpx/Vpr proteins for adaptation to its hostile cellular and species environments in the process of viral evolution.

Keywords: HIV-2, SIV, Vpx, Vpr, PPM

OPEN ACCESS

Edited by:
Akihide Ryo,
Yokohama City University, Japan

Reviewed by:
Mikako Fujita,
Kumamoto University, Japan
Mako Toyoda,
Kumamoto University, Japan

****Correspondence:***
Akio Adachi
adachi@tokushima-u.ac.jp
Masako Nomaguchi
nomaguchi@tokushima-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *21 April 2016*
Accepted: *20 July 2016*
Published: *03 August 2016*

Citation:
Sakai Y, Miyake A, Doi N, Sasada H, Miyazaki Y, Adachi A and Nomaguchi M (2016) Expression Profiles of Vpx/Vpr Proteins Are Co-related with the Primate Lentiviral Lineage. Front. Microbiol. 7:1211. doi: 10.3389/fmicb.2016.01211

INTRODUCTION

Human immunodeficiency virus types 1 and 2 (HIV-1 and HIV-2) are believed to be generated by extensive cross-species and/or intra-species transmissions of naturally occurring lentiviruses in African primates (Sharp and Hahn, 2011). To date, more than 40 primate species in Africa have been reported to harbor lentiviruses, structurally similar to HIV-1 and HIV-2 (Sharp and Hahn, 2011). Although the evolution and phylogeny of these viruses have been shown to be complicated (Sharp and Hahn, 2011; Shaw and Hunter, 2012; Swanstrom and Coffin, 2012), there are currently eight main lineages in HIV/simian immunodeficiency viruses (SIVs) (Peeters and Courgnaud, 2002; Gordon et al., 2005) (**Figure 1A**). The genomes of various HIV/SIVs individually contain a unique set of accessory genes designated *nef*, *vif*, *vpu*, *vpr* and *vpx* (**Figure 1B**). Accessory proteins encoded by these genes mainly function to inactivate host restriction factors, and thus optimize viral replication (Blanco-Melo et al., 2012; Harris et al., 2012; Malim and Bieniasz, 2012;

Simon et al., 2015). While all HIV/SIVs commonly have *nef*, *vif* and *vpr* genes, *vpu* and *vpx* genes are unique to some viral lineages. Upon directing at the *vpx*, *vpr*, and *vpu* genes, various HIV/SIVs can be grouped into three types (Fujita et al., 2010) (**Figure 1B**): prototype viruses with *vpr* only; HIV-1 type viruses carrying *vpr* and *vpu*; HIV-2 type viruses carrying *vpr* and *vpx*. Thus, *vpx* and *vpu* are unique to HIV-2 type and HIV-1 type viruses, respectively. Of note, Vpr and Vpx proteins show significant structural and functional similarities (Khamsri et al., 2006; Ayinde et al., 2010; Fujita et al., 2010). Among the major HIV/SIV lineages, viruses of the two groups, i.e., HIV-2/SIVsmm/stm/mac and SIVrcm/SIVdrl/mnd-2 (**Figure 1A**), have both Vpx and Vpr (**Figure 1B**).

Vpx is a virion-associated protein of 12–16 kDa, and exerts its function in the early stage of infection. Without functional Vpx, HIV-2 type viruses are unable or impeded to grow in natural target cells (Fujita et al., 2010). Recently, a cellular anti-viral factor SAMHD1 has been identified as target for Vpx (Hrecka et al., 2011; Laguette et al., 2011). However, a SAMHD1-independent mechanism(s) is still likely to exist (Fujita et al., 2012; Nomaguchi et al., 2014a; Schaller et al., 2014). From a structural point of view, although Vpx and Vpr are closely related and comprise three helices as described above, they are distinct from each other. Vpx has a zinc finger motif that stabilizes the helical structure (Schwefel et al., 2014), which is not present in Vpr. Notably, there is a well-conserved poly-proline motif (PPM), consisting of seven consecutive prolines, at the C-terminus of HIV-2 and SIVmac Vpx proteins (Miyake et al., 2014a). We previously showed that an HIV-2 mutant virus carrying multi-substitutional mutations in the PPM sequence did not grow at all in human macrophages and grew much more poorly than wild-type (WT) virus in a simian T-cell line, exactly like a Vpx-minus mutant (Fujita et al., 2008). Subsequent molecular studies demonstrated that PPM enhanced Vpx expression at a translation level, not influencing the stability of the protein (Miyake et al., 2014a,b). Our previous work also showed that HIV-1 and HIV-2 Vpr proteins were expressed at a much lower level relative to HIV-2 Vpx, and that the expression level of the two Vpr proteins was not enhanced significantly by simply adding the HIV-2 Vpx PPM sequence (Miyake et al., 2014a). Furthermore, despite a high overall homology of HIV-2 Vpx and SIVmac Vpx, their expression levels in transfected cells were significantly different (Miyake et al., 2014a).

In this report, we performed a linkage study between the Vpx expression profiles and viral phylogeny. Expression plasmids for a wide variety of Vpx proteins derived from diverse primate immunodeficiency viruses (**Figure 1**) were constructed, and monitored for their expression levels and PPM-dependency on the protein expression in transfected cells using SIVagm Vpr proteins as comparative controls. In parallel, phylogenetic trees based on Vpx/Vpr and Gag amino acid sequences were constructed to determine viral evolutionary relationships. The results obtained show that each viral lineage has its characteristic expression property, suggesting a link between the Vpx/Vpr expression pattern and viral evolutionary position.

MATERIALS AND METHODS

Virus Origins

Origins of SIVs are as follows (see also **Figure 1**). Prototype viruses: SIVagm (isolated from the African green monkey); SIVmnd-1 (mandrill); SIVlst (l'Hoest's monkey); SIVsun (sun-tailed monkey); SIVsyk (Sykes' monkey); SIVdeb (DeBrazza's monkey); SIVtal (talapoin monkey); SIVasc (red-tailed guenon); SIVcol (colobus monkey); SIVwrc (western red colobus); SIVolc (olive colobus); SIVkrc (Kibale red colobus). HIV-1 type viruses: SIVcpz (chimpanzee); SIVgor (gorilla); SIVgsn (greater spot-nosed monkey); SIVmon (mona monkey); SIVmus (mustached monkey); SIVden (Dent's monkey). HIV-2 type viruses: SIVsmm (sooty mangabey monkey); SIVmac (macaque monkey); SIVstm (stump-tailed macaque); SIVrcm (red-capped mangabey); SIVmnd-2 (mandrill); SIVdrl (drill monkey).

Plasmids

FLAG-tagged pEF-F expression plasmids for HIV-2 GL-AN Vpx (Genbank accession no., M30895), SIVmac239 Vpx (M33262), and their d7P (a complete deletion of seven consecutive prolines) mutants have been previously described (Miyake et al., 2014a). To generate new FLAG-tagged expression plasmids for Vpx/Vpr proteins in this study, *vpx* and *vpr* genes were synthesized (GenScript) and cloned into pEF-F as described above. New Vpx and Vpr proteins analyzed in this study are as follows (see also **Figure 2**): HIV-2 ALI Vpx (AF082339); HIV-2 EHO Vpx (U27200); HIV-2 Abt96 Vpx (AF208027); SIVsmm PGM53 Vpx (AF077017); SIVstm 37_16 Vpx (M83293); SIVsmm SL92B Vpx (AF334679); SIVmac 251BK28 Vpx (M19499); SIVrcm 02CM8081 Vpx (HM803689); SIVrcm GAB1 Vpx (AF382829); SIVrcm NG411 Vpx (AF349680); SIVdrl FAO Vpx (AY159321); SIVmnd-2 M14 Vpx (AF328295); SIVmnd-2 5440 Vpx (AY159322); SIVagm VER AGM3 Vpr (M30931); SIVagm TYO1 Vpr (DJ048201); SIVagm VER AGM155 Vpr (M29975); SIVagm VER 9063 Vpr (L40990); SIVagm GRI 677 Vpr (M66437). PPM-deletion mutants were constructed by the QuikChange site-directed mutagenesis kit (Agilent Technologies) or by overlap extension PCR using WT clones as templates.

Transfection

Human kidney 293T cells used for transfection experiments were cultured and maintained as previously described (Miyake et al., 2014a). For transfection, 4.0 μg of each plasmid DNA was introduced into 293T cells by 9.0 μl of Lipofectamine 2000 (Thermo Fisher Scientific).

Western Blotting

Western blot analysis of transfected cell lysates using anti-FLAG M2 antibody (Sigma) or anti-β-actin AC-15 antibody (Sigma) was conducted as described previously (Miyake et al., 2014a,b). Briefly, supernatants of cell lysates were prepared at 24 h post-transfection, and normalized for total protein

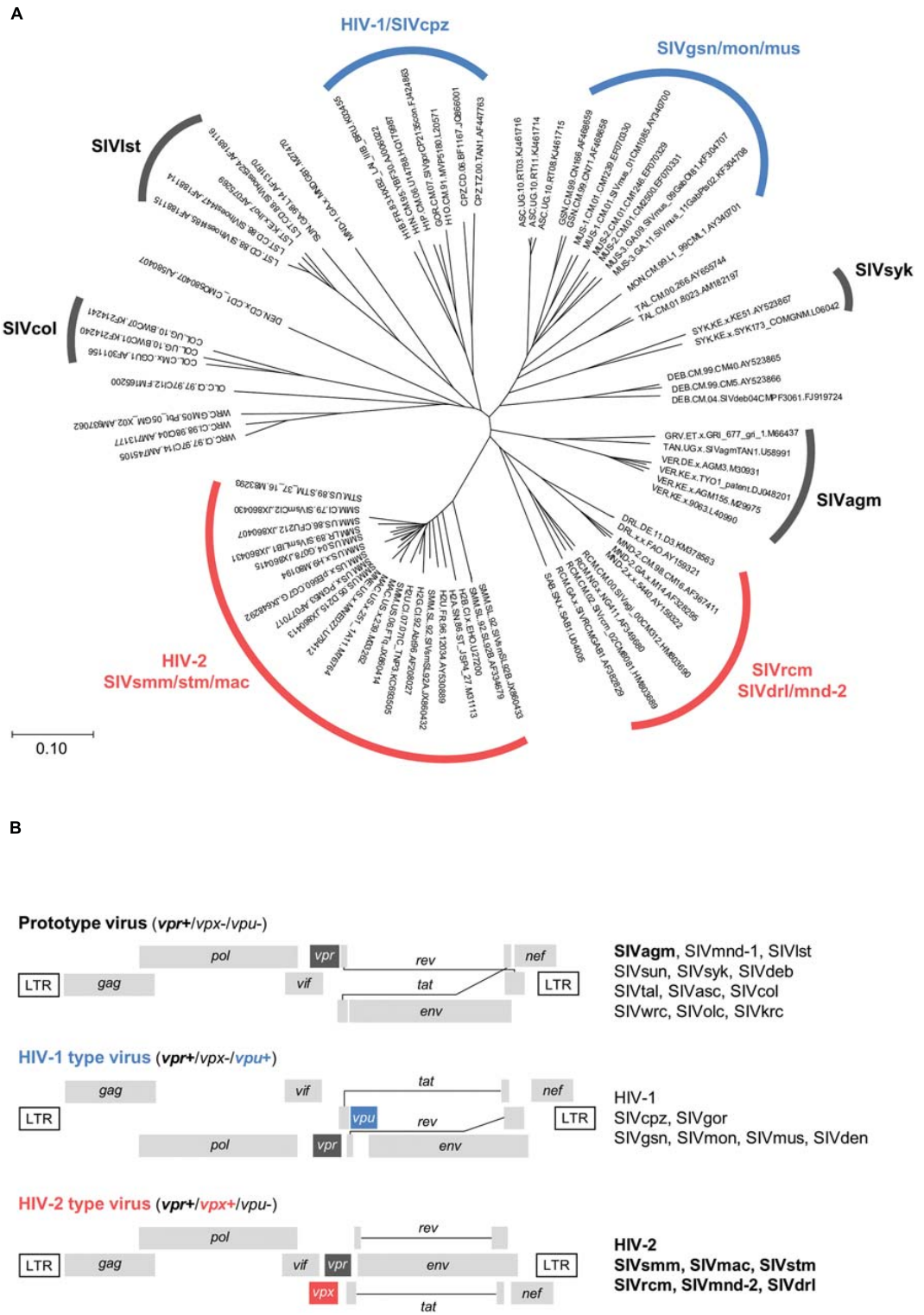


FIGURE 1 | Evolutionary relationship and genome structure of various HIV/SIVs. (A) Phylogeny of HIV/SIVs. The unrooted phylogenetic tree shown was inferred by the neighbor-joining method using amino acid sequences of the entire Gag polyprotein. Amino acid sequences in the HIV Sequence Compendium (http://www.hiv.lanl.gov) were used to generate the tree. Scale bar represents the genetic distance. Eight major viral lineages (Peeters and Courgnaud, 2002; Gordon et al., 2005) are marked as shown. Virus clones not yet classified into the lineage groups remain unmarked. Three genome types (Fujita et al., 2010) are indicated by black (prototype), blue (HIV-1 type), and red (HIV-2 type) letters/lines (see **B**). **(B)** Three types of the HIV/SIV genome organization. Genome structures are schematically shown. Letters in boldface type on the right show the lineages analyzed in this study. For virus designations, see Section "Materials and Methods."

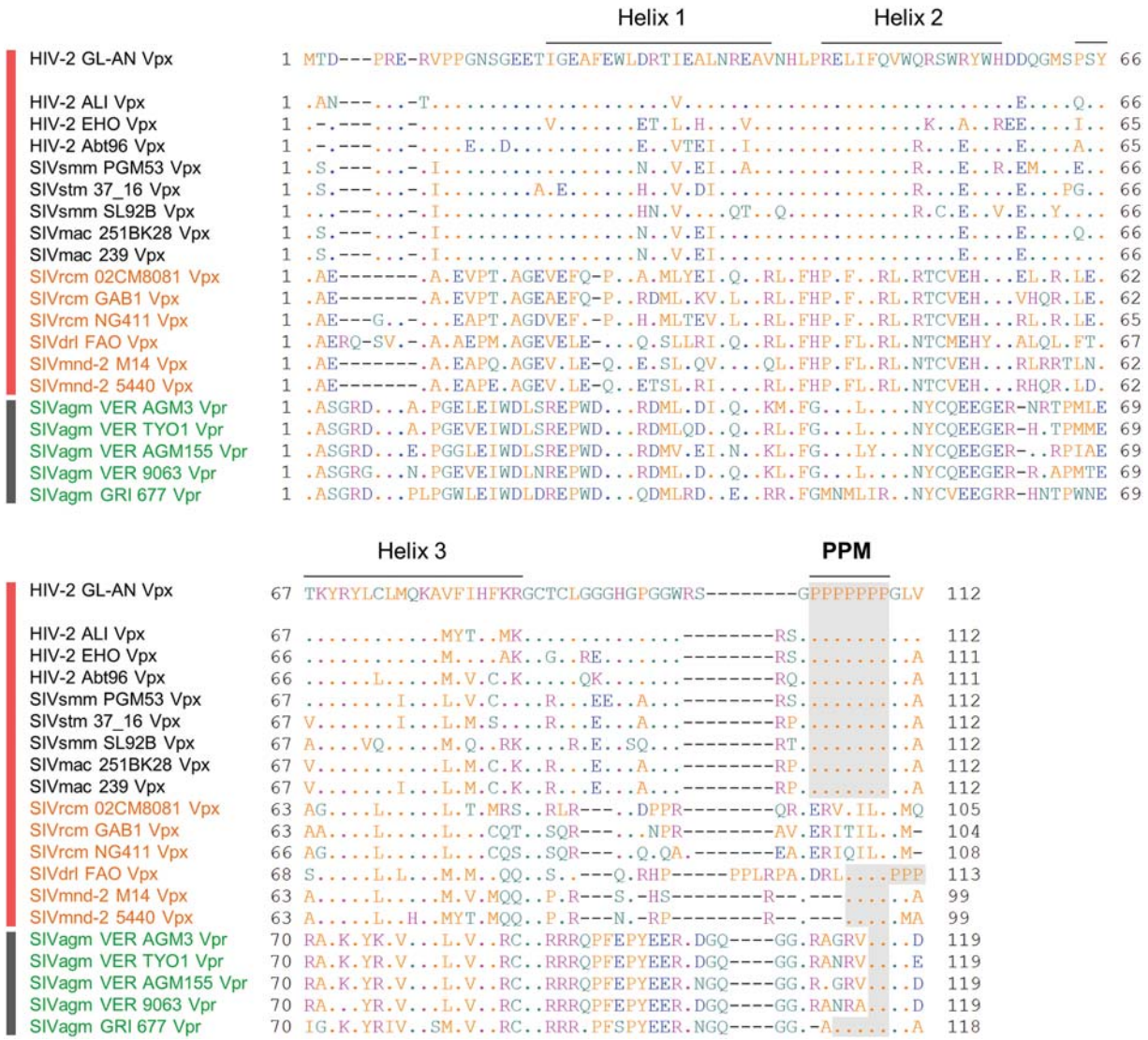


FIGURE 2 | Sequence alignment of various Vpx/Vpr proteins. Amino acid sequences of various Vpx/Vpr proteins analyzed in this study are shown. Various Vpx/Vpr proteins on the left derived from the HIV-2/SIVsmm/stm/mac, SIVrcm/SIVdrl/mnd-2, and SIVagm groups **(Figure 1A)** are represented by black, orange, and green letters, respectively. Viruses with HIV-2 type (HIV-2/SIVsmm/stm/mac and SIVrcm/SIVdrl/mnd-2 in **Figure 1B**) and prototype virus (SIVagm in **Figure 1B**) genomes are also indicated by vertical red and black bars, respectively. Sequences were obtained from the HIV sequence database at Los Alamos National Laboratory (http://www.hiv.lanl.gov) and aligned by Genetyx Ver. 11. Locations of three helices of HIV-2 GL-AN Vpx based on the references (Schaller et al., 2014; Schwefel et al., 2014) and a C-terminal PPM (two or more consecutive proline residues) region are indicated as shown.

amounts by the DC protein assay (Bio-Rad). Samples were then separated on any kD Mini-PROTEAN® TGX™ Precast Gels (Bio-Rad), and transferred onto PVDF membranes (Immobilon-P, Millipore). Immunoreactive viral and cellular proteins were visualized by chemiluminescence using Pierce Western Blotting Substrate Plus (Thermo Fisher Scientific). Experiments were performed at least three times, and representative results are shown. For quantification of the protein band intensities, a GS-800 calibrated densitometer and Quantity One software (Bio-Rad) were used. Mean values ± SD were obtained from at least three independent transfection experiments using HIV-2 GL-AN Vpx as a control.

Phylogenetic Analysis

Vpx, Vpr, and Gag proteins of HIV/SIVs were phylogenetically analyzed as previously described (Miyake et al., 2014a). Vpr and Gag proteins of SIVsyk were used as references. Amino acid sequences of Vpx, Vpr and Gag proteins deposited in the HIV sequence database at Los Alamos National Laboratory[1] were aligned by the CLUSTAL_X 2.0.11 program (Thompson et al., 1997; Jeanmougin et al., 1998). Phylogenetic trees were generated by the neighbor-joining method using CLUSTAL_X 2.0.11 program, and branch significance was

[1]http://www.hiv.lanl.gov

analyzed by bootstrap with 1000 replicates. Phylogenetic trees were visualized by GENETYX-Tree 2.2.2 program (Genetyx).

RESULTS

Sequence Features of Various Vpx Proteins

Through molecular and comparative analyses on Vpx proteins of HIV-2 GL-AN and SIVmac 239 clones, we previously identified two unique regions in Vpx that are important for Vpx expression in cells (Miyake et al., 2014a). One is PPM at the C-terminus and another is helix 1 in the N-terminal region. Deletion of the poly-proline stretch and/or introduction of multi-substitution mutations into the region greatly reduced the expression level of the two Vpx proteins in transfected cells relative to that of parental clones. Based on the observation that HIV-2 GL-AN and SIVmac 239 produce Vpx at a readily distinguishable level upon transfection, we made chimeric clones between the two to locate the determinant sequences that influence the expression level. The responsible region was mapped to four amino acids in the helix 1. We were interested in extending our previous study by monitoring the basal expression level and the PPM-dependency of Vpx proteins from diverse HIV/SIVs. We selected 15 viruses from various HIV/SIVs in the Los Alamos data base to represent the lineages carrying Vpx proteins (HIV-2/SIVsmm/stm/mac and SIVrcm/SIVdrl/mnd-2 groups in **Figure 1**) for analysis in this study (**Figure 2**). Three to five test clones that have no ambiguities in Vpx amino acid sequences were carefully chosen for each subgroup (HIV-2, SIVsmm/stm/mac, SIVrcm, and SIVdrl/mnd-2 in **Figure 1**) to minimize selection biases in the analysis. We focused on examining the Vpx expression here, but five SIVagm Vpr proteins were included because SIVagm Vpr was suggested as the origin of SIVsmm Vpx (Sharp et al., 1996). As readily observed in **Figure 2**, C-terminal PPM is well-conserved among the HIV-2/SIVsmm/stm/mac group and a SIVdrl/mnd-2 subgroup in the SIVrcm/SIVdrl/mnd-2 group (**Figure 1A**). However, no PPM is present at the corresponding region of Vpx proteins from another subgroup (SIVrcm) in the SIVrcm/SIVdrl/mnd-2 group (**Figure 1A**). Notably, there is a clear PPM consisting of five consecutive prolines in Vpr of an SIVagm strain (GRI 677) (**Figure 2**). Another point worth mentioning here is that Vpx/Vpr proteins are quite miscellaneous among the lineages, and there are scattered amino acid differences even in the helix region of the proteins from the same viral lineage (**Figure 2**). This is true for the helix 1 region of Vpx proteins in the HIV-2/SIVsmm/stm/mac group (**Figure 2**).

Expression Profiles for Various Vpx/Vpr Proteins

We assumed, based on empirical knowledge (Khamsri et al., 2006; Fujita et al., 2008; Miyake et al., 2014a), that some Vpx/Vpr proteins derived from diverse HIV/SIVs would be produced at a very low level upon transfection. In order to perform a systemic quantitative analysis on the expression of these proteins, especially to detect and compare a minimal level expression, a highly efficient transfection method generating highly reproducible results is required. Although we employed the calcium-phosphate co-precipitation method in previous studies (Miyake et al., 2014a,b), we selected here to use the lipofection method instead because of its better reproducibility. First, we re-evaluated, by this new method, the effect of PPM-deletion on the Vpx expression level of HIV-2 in a quantitative manner. Parental and mutant clones were transfected into 293T cells, and 24 h later, sample cell lysates were prepared for Western blot analysis. As is clear in **Figure 3**, the d7P mutant expressed mutant Vpx protein at a level between an eighth and a sixteenth relative to that by WT clone. The d7P mutant protein was readily detected in both experiments, and the PPM-dependent expression of HIV-2 Vpx was evident in both experiments.

We then comparatively and quantitatively assessed the expression levels of numerous Vpx and Vpr proteins derived from a variety of HIV/SIVs. Eighteen new expression plasmids with a FLAG-tag were constructed (**Figure 2**), and examined for their expression in 293T cells following transfection. HIV-2 (GL-AN) Vpx was used as a control throughout the experiments. **Figure 4** shows the representative results obtained by this all-inclusive monitoring. As predicted, the Vpx/Vpr expression levels, by viral clones belonging to the different groups (**Figures 4A–E** correspond to the HIV-2, SIVsmm/stm/mac, SIVrcm, SIVdrl/mnd-2, and SIVagm groups, respectively), significantly or clearly varied. In addition, Vpx and Vpr proteins were produced at a different level even by viruses within the same group, and no clear group-specificity with respect to the Vpx/Vpr expression level in cells was observed. Of note, small differences were observed for Vpx proteins from the SIVdrl/mnd-2 group (**Figure 4D**). Remarkably, some SIVagm Vpr proteins were expressed at an extremely low level

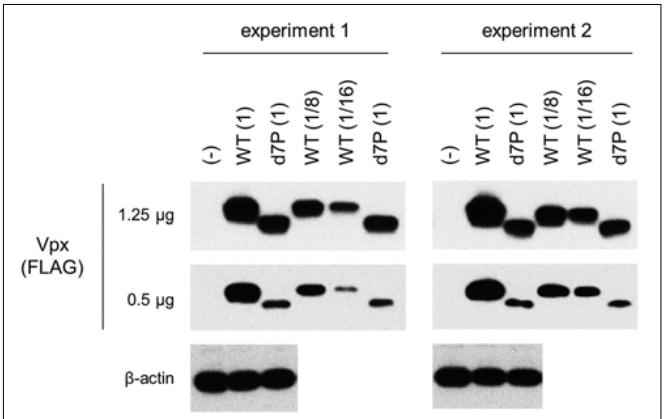


FIGURE 3 | Expression levels of HIV-2 GL-AN Vpx and its d7P mutant. Expression levels of wild-type (WT) and mutant Vpx proteins in transfected 293T cells were examined by Western blotting analysis. Results of two independent experiments are shown. Relative protein amounts used for blotting are indicated in parentheses. As WT samples, an eighth and a sixteenth amounts were also used for comparative quantification purpose. (–), pEF1/*Myc*-HisA (empty vector); WT, pEF-Fvpx; d7P, pEF-Fvpx/d7P mutant (Miyake et al., 2014a,b).

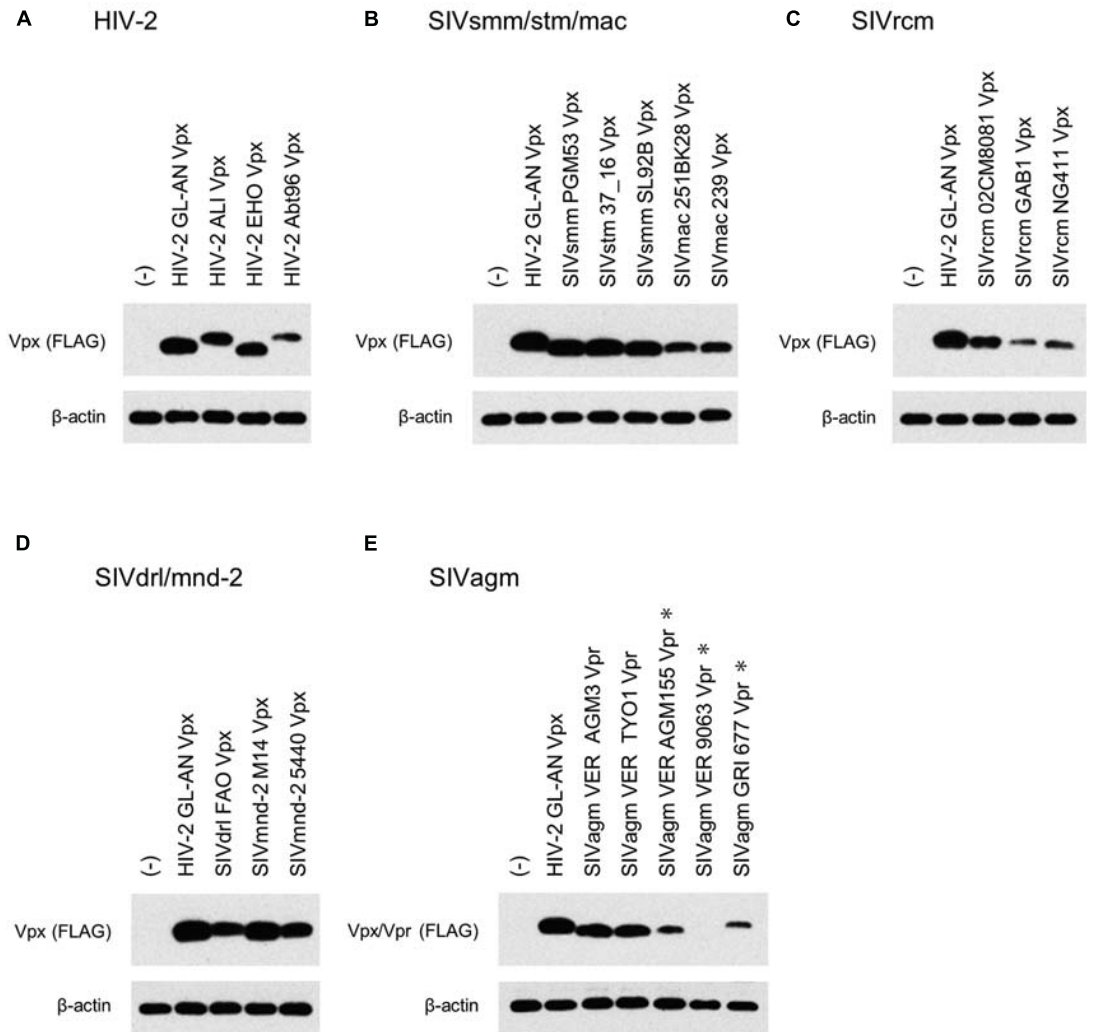


FIGURE 4 | Expression of Vpx/Vpr proteins from various HIV/SIVs. Expression of various Vpx/Vpr proteins **(A–E)** in transfected 293T cells as monitored by Western blotting analysis is shown. Viral groups **(Figure 1A)** are indicated. *, 10-fold more protein amounts were used to detect Vpr, but same protein amounts with those for the other four lanes were used to detect β-actin here. (–), pEF1/*Myc*-HisA (empty vector).

(**Figure 4E**). To better substantiate the results in **Figure 4**, we quantified the expression levels by densitometric monitoring of the Vpx/Vpr band intensities, and calculated the levels relative to that by the control HIV-2 GL-AN. As shown in **Figure 5**, the expression levels could be categorized into four groups: high (>70% relative to GL-AN), medium (30–70%), low (<30%), and ultra-low (minimum expression). Overall, these results quantitatively confirmed that the expression levels of Vpx/Vpr vary considerably among primate lentiviruses.

In order to determine the PPM dependency of Vpx/Vpr expression by clones containing PPM (two or more consecutive prolines), we constructed PPM-deletion mutants from various virus species (**Figure 2**), and examined their expressions relative to parental clones. **Figure 6** shows the results obtained for each viral group: HIV-2/SIVsmm/stm/mac in panel A; SIVdrl/mnd-2 in B; SIVagm in C. As clearly observed, most Vpx/Vpr proteins exhibited the PPM-dependency except for those from viruses in the SIVdrl/mnd-2 group (**Figure 6B**). Unexpectedly, the expression levels of Vpr proteins derived from the two virus strains in the SIVagm group were enhanced by PPM-deletion (**Figure 6C**), in a sharp contrast to the results for the others. In total, the results in **Figure 6** revealed that the PPM-dependent expression of Vpx, including SIVagm Vpr, is a conserved feature among most HIV-2/SIVs in the transfected 293T cells.

Phylogenetic Study

We constructed a phylogenetic tree of various Vpx/Vpr proteins to determine whether the expression profiles of viral Vpx/Vpr proteins presented so far could be related to viral evolutional positions (**Figure 7**). As is recognizable in this figure, evolutional group-dependent properties for the Vpx/Vpr expression pattern became clear. While members in the SIVdrl/mnd-2 group expressed a relatively high level of Vpx without PPM-dependence, clones in the SIVrcm group do not have the

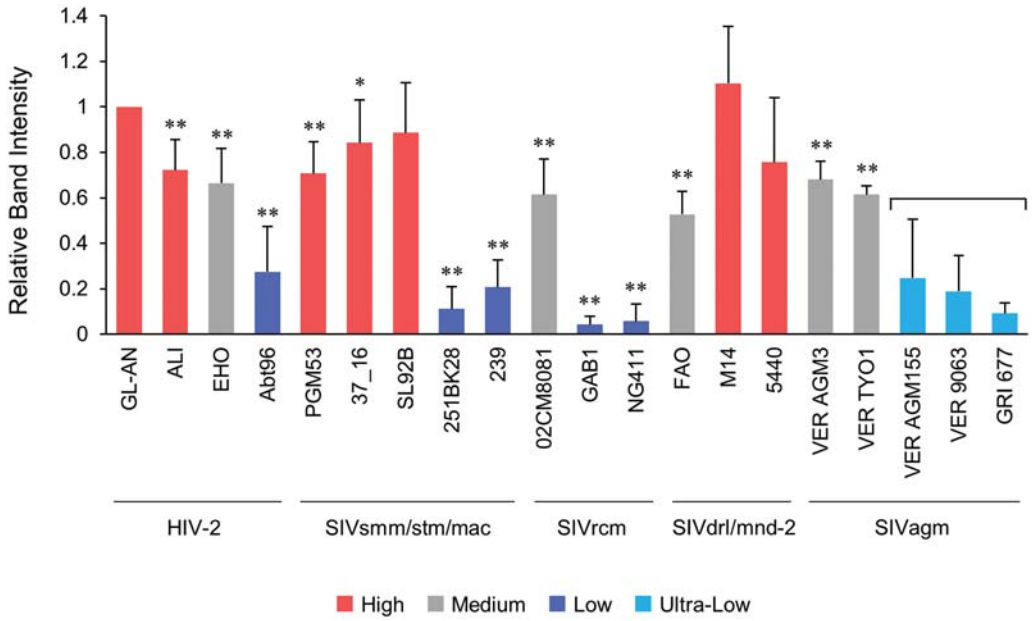


FIGURE 5 | Quantification of the Vpx/Vpr expression levels. Following visualization of blotted Vpx/Vpr proteins, the band intensities were determined by a densitometer, and the expression levels relative to that of HIV-2 GL-AN Vpx were calculated. Based on the mean intensity values, the expression levels were grouped into High (>70% relative to GL-AN), Medium (30–70%), Low (<30%), and Ultra-Low (undetectable under the standard condition). For Ultra-Low group, 10-fold more protein amounts were used to perform Western blotting analysis. Viral groups **(Figure 1A)** are indicated under the lines. Mean values ± SD from at least three independent experiments are shown. Significance relative to HIV-2 GL-AN as calculated by the Student *t*-test is shown (*$P < 0.05$; **$P < 0.01$).

PPM itself. Although the Vpx expression levels observed for the SIVrcm clones were significantly different, this could be a subgroup difference. In the large HIV-2/SIVsmm/stm/mac group, all clones share the C-terminal seven consecutive prolines, and all the clones examined exhibited the PPM-dependency for their Vpx expression. Considering the branching to subgroups, the distinct expression levels from (L) to (H) in this large group may be evolutionarily explainable. As observed in the phylogenetic tree (**Figure 7**), some viruses in one subgroup may have acquired or lost their properties in the course of adaptation and evolution to form the other subgroups (for example, see SIVsmm SL92B, HIV-2 EHO, and HIV-2 Abt96). Finally, two strains (VER 9063 and VER AGM155) in the SIVagm group that show unique PPM-dependency were positioned relatively close to each other within the group. In total, our results here suggest a possible link between the expression profiles (the basal expression level and PPM-dependency in the transfected 293T cells) of Vpx/Vpr proteins and the primate lentiviral evolutional positions.

The phylogenetic tree described above was based on Vpx/Vpr amino acid sequences. Therefore, it was possible that our experimental data on various HIV/SIVs (**Figures 4–6**) simply reflected the amino acid sequence similarity *per se*, not being indicative of accurate evolutionary history. To exclude this possibility, we constructed a phylogenetic tree based on major structure protein Gag, and compared the two phylogenetic trees. As is clear in **Figure 8**, the branching pattern in Gag-tree was generally consistent with that in Vpx/Vpr-tree (**Figure 7**), ruling out the above possibility.

DISCUSSION

In this study, we constructed numerous expression plasmids for diverse Vpx/Vpr proteins derived from HIV-2 type and prototype viruses (**Figure 1**), and examined their basal and PPM-dependent expression phenotypes in cells (**Figures 4–6**). We also constructed phylogenetic trees based on viral Vpx/Vpr and Gag proteins (**Figures 7** and **8**). Basal expression levels of Vpx/Vpr proteins were found to be highly variable, but appeared to be evolutional group-dependent (**Figures 7** and **8**). PPM-dependent or -independent expression of Vpx/Vpr was demonstrated to be evolutional group-specific (**Figures 7** and **8**). Taken together, it is not unreasonable to conclude that various HIV/SIVs may have acquired characteristic abilities to express Vpx/Vpr proteins to meet with their surrounding circumstances. This is the first report that suggests the evolutional significance of Vpx/Vpr expression patterns.

Two strains in the SIVsmm/stm/mac group analyzed here (SIVmac 239 and 251BK28) expressed Vpx at a low level, whereas closely related viruses of this group all produced Vpx at a high level (SIVsmm PGM53 and SL92B, and SIVstm 37_16) (**Figure 5**). Probably, the low level phenotype was newly acquired by SIVmac when branched from SIVsmm. Similarly, viral strains GAB1 and NG411 in the SIVrcm group expressed Vpx at a low level, but another member 02CM8081 did at a medium level (**Figure 5**). Because GAB1 and NG411 cluster together to form a separate subgroup from the 02CM8081 group on the phylogenetic tree (**Figure 7**), one can postulate that different expression phenotypes were independently acquired. Viruses in

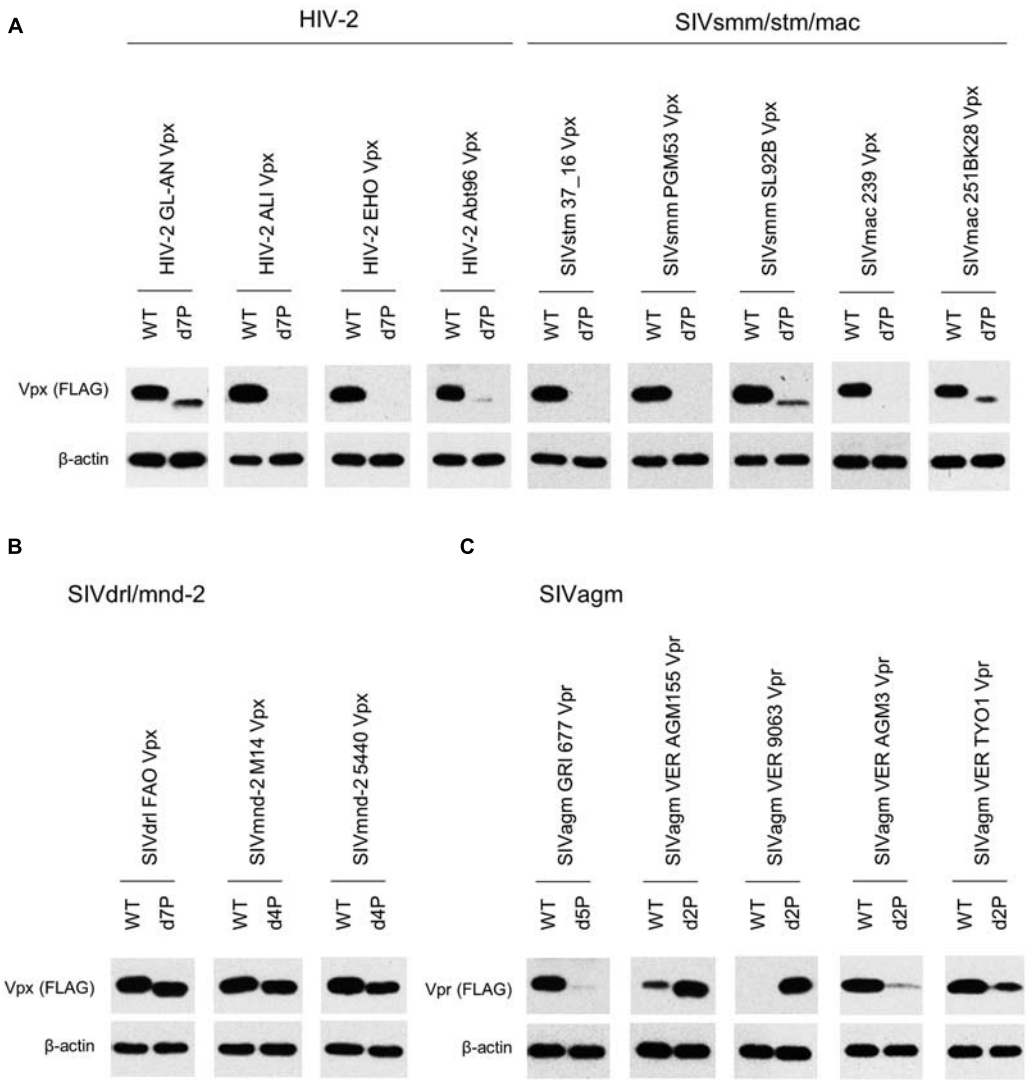


FIGURE 6 | Expression of PPM-deleted Vpx/Vpr mutant proteins from various HIV/SIVs. Expression of PPM-deleted Vpx/Vpr proteins **(A–C)** in transfected 293T cells as monitored by Western blotting analysis is shown. Viral groups **(Figure 1A)** are indicated. Each deletion mutant (d2P, d4P, d5P, and d7P) lacks the entire PPM sequence **(Figure 2)**. Amounts of sample cell lysates, antibody concentrations, and exposure time were adjusted individually for appropriate band intensity as shown.

the SIVrcm group are particularly interesting, for only they carry Vpx without PPM. Highly divergent sequences of SIVrcm Vpx (Beer et al., 2001) might be associated with this unique feature. Whereas SIVsmm Vpx has PPM of seven-consecutive prolines not found in SIVrcm Vpx, SIVrcm and SIVsmm were suggested to gain their Vpx proteins before their divergence (Etienne et al., 2013). Therefore, it is intriguing to elucidate why and how these two viruses came to encode distinct Vpx proteins with/without PPM.

The SIVdrl/SIVmnd-2 group exceptionally lacks PPM-dependency for its Vpx expression (**Figure 6**). Viruses in this group have a relatively long PPM (four or seven consecutive prolines) (**Figure 2**) and expressed Vpx at a medium or high level (**Figure 5**). SIVdrl and SIVmnd-2 have mosaic genome structures and were suggested to have arisen from a recombinational event(s) between SIVrcm with *vpx* and SIVmnd-1 without *vpx* (**Figure 1**) (Takemura and Hayami, 2004). Therefore, Vpx proteins of SIVdrl and SIVmnd-2 are predicted to be originated from an ancient SIVrcm Vpx. Because SIVrcm Vpx lacks PPM (**Figure 2**), one needs to assume that SIVdrl and SIVmnd-2 newly acquired Vpx with PPM in the evolution process. Or, it could be supposed that there was PPM in the ancient SIVrcm Vpx, and that the PPM was subsequently lost from the Vpx (after divergence of SIVdrl and SIVmnd-2 from SIVrcm).

Remarkably, two SIVagm Vpr proteins (VER 9063 and VER AGM155) were expressed at a faint level in cells, and have a di-proline motif that suppresses protein expression (**Figure 7**). While two other SIVagm Vpr proteins (VER AGM3 and

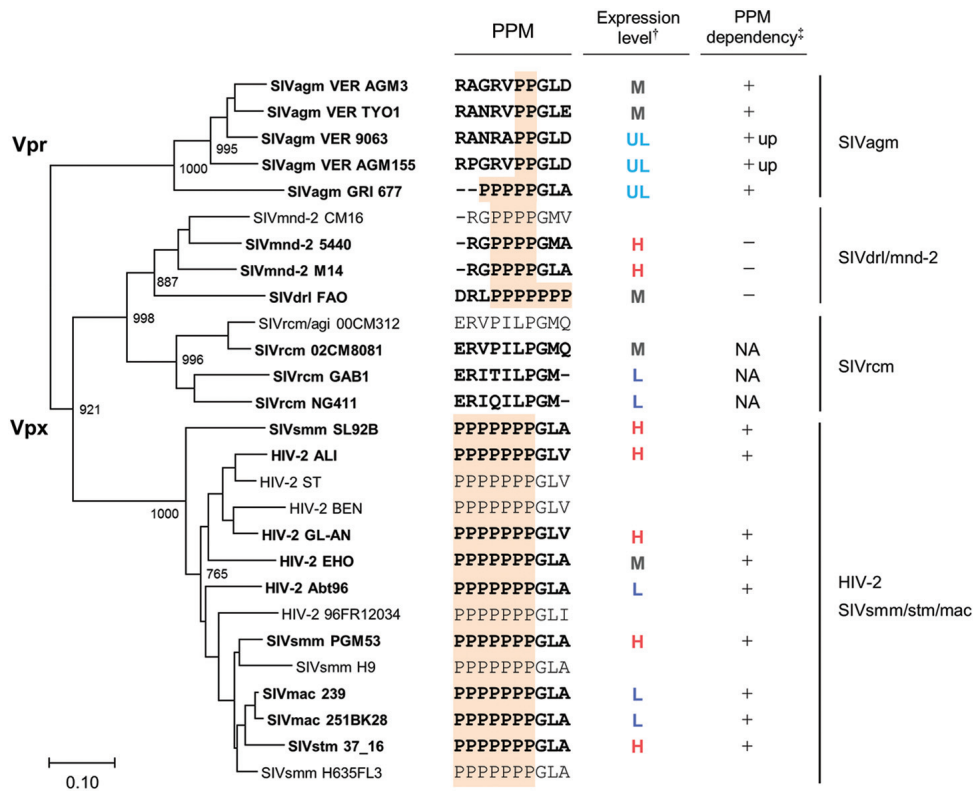


FIGURE 7 | Phylogeny and expression profiles of Vpx/Vpr proteins. Phylogenetic tree based on Vpx/Vpr proteins and their expression profiles in cells are shown in parallel (see **Figure 1A** for viral groups on the right). Scale bar represents the genetic distance. Branches were calculated from 1000 bootstrap replicates, and the bootstrap values are labeled on the major branches. Viral strains experimentally analyzed in this study (20 strains) are highlighted by boldface type letters. C-terminal regions of Vpx/Vpr proteins including PPM sequence (highlighted) and the summarized expression profiles **(Figures 4–6)** are shown in the middle. †: H, high; M, medium; L, low; UL, ultra-low (see **Figure 5**). ‡: +, decreased expression by PPM-deletion; +up, increased expression by PPM-deletion; −, no clear effect by PPM-deletion; NA, not applicable (no PPM).

VER TYO1) exhibited an ordinary phenotype (medium level expression and PPM-dependency), the fifth one (GRI 677) showed a unique character, i.e., ultra-low level expression and PPM-dependency, in the group (**Figure 7**). Viruses in the SIVagm group were reported to have a greater genetic diversity compared to other primate lentiviruses (Johnson et al., 1990), in agreement with a complex phenotype observed in this study as described above (**Figure 7**). SIVagm *vpr* gene was initially categorized as *vpx* due to its apparent sequence similarity to HIV-2 and SIVsmm *vpx* genes. However, it was subsequently reclassified as *vpr*, and also was proposed as the origin of SIVsmm *vpx* (Sharp et al., 1996). The *vpx* gene is found only in HIV-2 type viruses (**Figure 1**) and all these viruses were presumably originated from SIVagm. Thus, it can be assumed that SIVagm Vpr with its own PPM sequence evolved into various Vpx proteins. Recently, a similar hypothesis was presented for the SAMHD1 counteraction by Vpx and Vpr proteins (Lim et al., 2012). In the report, phylogeny-based observation has suggested that the SAMHD1 antagonism by *vpr* preceded the appearance of *vpx*.

A high divergence in the expression profiles observed for Vpx/Vpr proteins is virologically important. Various HIV/SIVs appeared to express Vpx/Vpr proteins at levels of an unexpectedly wide range (more than 100-fold difference) (**Figures 4, 5**, and **7**). We previously reported that the helix 1 (four amino acids) was a determinant for the Vpx expression level (Miyake et al., 2014a). However, taking account of considerably variable amino acid sequences in the helix 1 (**Figure 2**) of Vpx proteins with low and high expression phenotypes (**Figures 4, 5**, and **7**), we cannot simply conclude now that the helix 1 is responsible for determining the expression levels. In addition, some Vpx/Vpr proteins are PPM-dependent and others are PPM-independent for their expressions (**Figures 6** and **7**). Furthermore, there is no PPM in some Vpx proteins derived from one viral group. Considering that Vpx/Vpr proteins are very likely to be important for HIV/SIV replication, persistence and/or transmission in natural primate hosts, the heterogeneity with respect to the expression level of a viral protein is remarkable. However, of note, the phylogeny of various HIV/SIVs generally shows a cluster pattern similar with that of the natural hosts (Perelman et al., 2011), supporting the notion that HIV/SIVs have acquired various expression profiles of Vpx/Vpr in the course of virus diversification. Although presently unknown, biological and molecular bases/reasons for the observations described above must exist. Our work reported here would evoke studies

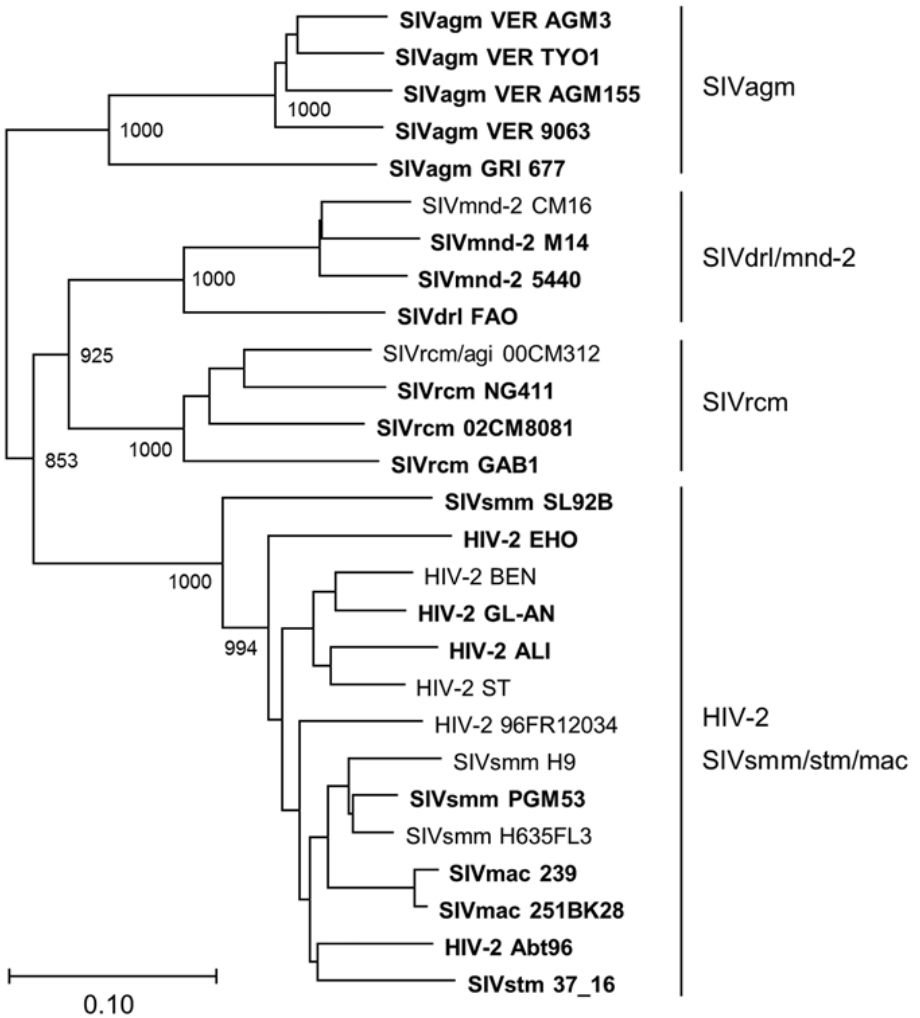


FIGURE 8 | Phylogeny of Gag proteins. Phylogenetic tree based on Gag proteins is shown. Viral groups **(Figure 1A)** are indicated on the right. Viral strains experimentally analyzed in this study (20 strains) are highlighted by boldface type letters. Scale bar indicates the genetic distance. Branches were calculated from 1000 bootstrap replicates, and the bootstrap values are labeled on the major branches.

to determine the processes and underlying molecular bases (transcription, translation, stability of Vpx/Vpr, cytotoxicity of Vpx/Vpr, responsible sequence or structure of Vpx/Vpr, relevant cellular factors, and so on) for the observed heterogeneity among various HIV/SIVs. In this regard, of note, our recent studies have demonstrated that HIV-1 can adapt itself to various APOBEC3G environments by regulating the Vif expression level (Nomaguchi et al., 2014b, 2016).

Although we demonstrate here that Vpx/Vpr expression profiles are potentially linked to the phylogeny of various HIV/SIVs, our research system used was rather artificial, and how the Vpx/Vpr proteins contribute to the biology of HIV/SIVs in host individuals and populations remains to be elucidated. Further studies utilizing infectious intact proviral clones derived from various HIV/SIVs and natural target cells from various primate hosts are required to reveal the biological role for Vpx/Vpr in the process of viral adaptation and evolution.

AUTHOR CONTRIBUTIONS

YS: acquisition, analysis, and interpretation of data for the work; drafting the work; final approval of the manuscript. AM: acquisition, analysis, and interpretation of data for the work; final approval of the manuscript. ND: acquisition, analysis, and interpretation of data for the work; final approval of the manuscript. HS: acquisition, analysis, and interpretation of data for the work; final approval of the manuscript. YM: acquisition, analysis, and interpretation of data for the work; final approval of the manuscript. AA: design of the work; analysis, and interpretation of data for the work; drafting the work; final approval of the manuscript agreement to be accountable for all aspects of the work. MN: design of the work; analysis, and interpretation of data for the work; drafting the work; final approval of the manuscript agreement to be accountable for all aspects of the work.

FUNDING

This study is supported in part by a Grant-in-Aid for Scientific Research (B) (26293104) to AA from the Japan Society for the Promotion of Science.

ACKNOWLEDGMENT

We thank Ms. Kazuko Yoshida (Department of Microbiology, Tokushima University Graduate School of Medical Science, Tokushima 770-8503, Japan) for her editorial assistance.

REFERENCES

Ayinde, D., Maudet, C., Transy, C., and Margottin-Goguet, F. (2010). Limelight on two HIV/SIV accessory proteins in macrophage infection: is Vpx overshadowing Vpr? *Retrovirology* 7:35. doi: 10.1186/1742-4690-7-35

Beer, B. E., Foley, B. T., Kuiken, C. L., Tooze, Z., Goeken, R. M., Brown, C. R., et al. (2001). Characterization of novel simian immunodeficiency viruses from red-capped mangabeys from Nigeria (SIVrcmNG409 and -NG411). *J. Virol.* 75, 12014–12027. doi: 10.1128/JVI.75.24.12014-12027.2001

Blanco-Melo, D., Venkatesh, S., and Bieniasz, P. D. (2012). Intrinsic cellular defenses against human immunodeficiency viruses. *Immunity* 37, 399–411. doi: 10.1016/j.immuni.2012.08.013

Etienne, L., Hahn, B. H., Sharp, P. M., Matsen, F. A., and Emerman, M. (2013). Gene loss and adaptation to hominids underlie the ancient origin of HIV-1. *Cell Host Microbe* 14, 85–92. doi: 10.1016/j.chom.2013.06.002

Fujita, M., Nomaguchi, M., Adachi, A., and Otsuka, M. (2012). SAMHD1-dependent and -independent functions of HIV-2/SIV Vpx protein. *Front. Microbiol.* 3:297. doi: 10.3389/fmicb.2012.00297

Fujita, M., Otsuka, M., Nomaguchi, M., and Adachi, A. (2008). Functional region mapping of HIV-2 Vpx protein. *Microbes Infect.* 10, 1387–1392. doi: 10.1016/j.micinf.2008.08.005

Fujita, M., Otsuka, M., Nomaguchi, M., and Adachi, A. (2010). Multifaceted activity of HIV Vpr/Vpx proteins: the current view of their virological functions. *Rev. Med. Virol.* 20, 68–76. doi: 10.1002/rmv.636

Gordon, S., Pandrea, I., Dunham, R., Apetrei, C., and Silvestri, G. (2005). "The call of the wild: what can be learned from studies of SIV infection of natural hosts?" in *HIV Sequence Compendium,* eds T. Thomas Leitner, B. Foley, B. Hahn, P. Marx, F. McCutchan, J. Mellors, et al. (Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory).

Harris, R. S., Hultquist, J. F., and Evans, D. T. (2012). The restriction factors of human immunodeficiency virus. *J. Biol. Chem.* 287, 40875–40883. doi: 10.1074/jbc.R112.416925

Hrecka, K., Hao, C., Gierszewska, M., Swanson, S. K., Kesik-Brodacka, M., Srivastava, S., et al. (2011). Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474, 658–661. doi: 10.1038/nature10195

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., and Gibson, T. J. (1998). Multiple sequence alignment with CLUSTAL_X. *Trends Biochem. Sci.* 23, 403–405. doi: 10.1016/S0968-0004(98)01285-7

Johnson, P. R., Fomsgaard, A., Allan, J., Gravell, M., London, W. T., Olmsted, R. A., et al. (1990). Simian immunodeficiency viruses from African green monkeys display unusual genetic diversity. *J. Virol.* 64, 1086–1092.

Khamsri, B., Murao, F., Yoshida, A., Sakurai, A., Uchiyama, T., Shirai, H., et al. (2006). Comparative study on the structure and cytopathogenic activity of HIV Vpr/Vpx proteins. *Microbes Infect.* 8, 10–15. doi: 10.1016/j.micinf.2005.05.020

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Ségéral, E., et al. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474, 654–657. doi: 10.1038/nature10117

Lim, E. S., Fregoso, O. I., McCoy, C. O., Matsen, F. A., Malik, H. S., and Emerman, M. (2012). The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. *Cell Host Microbe* 11, 194–204. doi: 10.1016/j.chom.2012.01.004

Malim, M. H., and Bieniasz, P. D. (2012). HIV restriction factors and mechanisms of evasion. *Cold Spring Harb. Perspect. Med.* 2:a006940. doi: 10.1101/cshperspect.a006940

Miyake, A., Fujita, M., Fujino, H., Koga, R., Kawamura, S., Otsuka, M., et al. (2014a). Poly-proline motif in HIV-2 Vpx is critical for its efficient translation. *J. Gen. Virol.* 95, 179–189. doi: 10.1099/vir.0.057364-0

Miyake, A., Miyazaki, Y., Fujita, M., Nomaguchi, M., and Adachi, A. (2014b). Role of poly-proline motif in HIV-2 Vpx expression. *Front. Microbiol.* 5:24. doi: 10.3389/fmicb.2014.00024

Nomaguchi, M., Doi, N., and Adachi, A. (2014a). Virological characterization of HIV-2 *vpx* gene mutants in various cell systems. *Microbes Infect.* 16, 695–701. doi: 10.1016/j.micinf.2014.06.004

Nomaguchi, M., Doi, N., Sakai, Y., Ode, H., Iwatani, Y., Ueno, T., et al. (2016). Natural single-nucleotide variations in the HIV-1 genomic SA1prox region can alter viral replication ability by regulating Vif expression levels. *J. Virol.* 90, 4563–4578. doi: 10.1128/JVI.02939-15

Nomaguchi, M., Miyake, A., Doi, N., Fujiwara, S., Miyazaki, M., Tsunetsugu-Yokota, Y., et al. (2014b). Natural single-nucleotide polymorphisms in the 3' region of the HIV-1 pol gene modulate viral replication ability. *J. Virol.* 88, 4145–4160. doi: 10.1128/JVI.01859-13

Peeters, M., and Courgnaud, V. (2002). "Overview of primate lentiviruses and their evolution in non-human primates in Africa," in *HIV Sequence Compendium,* eds C. Kuiken, B. Foley, E. Freed, B. Hahn, B. Korber, P. A. Marx, et al. (Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory).

Perelman, P., Johnson, W. E., Roos, C., Seuánez, H. N., Horvath, J. E., Moreira, M. A. M., et al. (2011). A molecular phylogeny of living primates. *PLoS Genet.* 7:e1001342. doi: 10.1371/journal.pgen.1001342

Schaller, T., Bauby, H., Hué, S., Malim, M. H., and Goujon, C. (2014). New insights into an X-traordinary viral protein. *Front. Microbiol.* 5:126. doi: 10.3389/fmicb.2014.00126

Schwefel, D., Groom, H. C., Boucherit, V. C., Christodoulou, E., Walker, P. A., Stoye, J. P., et al. (2014). Structural basis of lentiviral subversion of a cellular protein degradation pathway. *Nature* 505, 234–238. doi: 10.1038/nature12815

Sharp, P. M., Bailes, E., Stevenson, M., Emerman, M., and Hahn, B. H. (1996). Gene acquisition in HIV and SIV. *Nature* 383, 586–587. doi: 10.1038/383586a0

Sharp, P. M., and Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harb. Perspect. Med.* 1:a006841. doi: 10.1101/cshperspect.a006841

Shaw, G. M., and Hunter, E. (2012). HIV transmission. *Cold Spring Harb. Perspect. Med.* 2:a006965. doi: 10.1101/cshperspect.a006965

Simon, V., Bloch, N., and Landau, N. R. (2015). Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nat. Immunol.* 16, 546–553. doi: 10.1038/ni.3156

Swanstrom, R., and Coffin, J. (2012). HIV-1 pathogenesis: the virus. *Cold Spring Harb. Perspect. Med.* 2:a007443. doi: 10.1101/cshperspect.a007443

Takemura, T., and Hayami, M. (2004). Phylogenetic analysis of SIV derived from mandrill and drill. *Front. Biosci.* 9:513–520. doi: 10.2741/1242

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Sakai, Miyake, Doi, Sasada, Miyazaki, Adachi and Nomaguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

PERSPECTIVE
published: 18 October 2016
doi: 10.3389/fmicb.2016.01655

Phylogenetic Insights into the Functional Relationship between Primate Lentiviral Reverse Transcriptase and Accessory Proteins Vpx/Vpr

Yosuke Sakai[1], Naoya Doi[1], Yasuyuki Miyazaki[2], Akio Adachi[1] and Masako Nomaguchi[1]**

[1] Department of Microbiology, Tokushima University Graduate School of Medical Science, Tokushima, Japan, [2] Department of Microbiology and Cell Biology, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

The efficiency of reverse transcription to synthesize viral DNA in infected cells greatly influences replication kinetics of retroviruses. However, viral replication in non-dividing cells such as resting T cells and terminally differentiated macrophages is potently and kinetically restricted by a host antiviral factor designated SAMHD1 (sterile alpha motif and HD-domain containing protein 1). SAMHD1 reduces cellular deoxynucleoside triphosphate (dNTP) pools and affects viral reverse transcription step. Human immunodeficiency virus type 2 (HIV-2) and some simian immunodeficiency viruses (SIVs) have Vpx or Vpr to efficiently degrade SAMHD1. Interestingly, the reverse transcriptase (RT) derived from HIV-1 that encodes no anti-SAMHD1 proteins has been previously demonstrated to uniquely exhibit a high enzymatic activity. It is thus not irrational to assume that some viruses may have acquired or lost the specific RT property to better adapt themselves to the low dNTP environments confronted in non-dividing cells. This adaptation process may probably be correlated with the SAMHD1-antagonizing ability by viruses. In this report, we asked whether such adaptive events can be inferable from Vpx/Vpr and RT phylogenetic trees overlaid with SAMHD1-degrading capacity of Vpx/Vpr and with kinetic characteristics of RT. Resultant two trees showed substantially similar clustering patterns, and therefore suggested that the properties of RT and Vpx/Vpr can be linked. In other words, HIV/SIVs may possess their own RT proteins to adequately react to various dNTP circumstances in target cells.

Keywords: HIV, SIV, Vpx, Vpr, RT, SAMHD1, dNTP

OPEN ACCESS

Edited by:
Akihide Ryo,
Yokohama City University, Japan

Reviewed by:
Hiroaki Takeuchi,
Tokyo Medical and Dental University, Japan
Ayumi Kudoh,
Yokohama City University, Japan

****Correspondence:***
Akio Adachi
adachi@tokushima-u.ac.jp
Masako Nomaguchi
nomaguchi@tokushima-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *23 August 2016*
Accepted: *04 October 2016*
Published: *18 October 2016*

Citation:
Sakai Y, Doi N, Miyazaki Y, Adachi A and Nomaguchi M (2016) Phylogenetic Insights into the Functional Relationship between Primate Lentiviral Reverse Transcriptase and Accessory Proteins Vpx/Vpr. Front. Microbiol. 7:1655. doi: 10.3389/fmicb.2016.01655

INTRODUCTION

Accessory proteins of various human immunodeficiency viruses/simian immunodeficiency viruses (HIV/SIVs) are believed to be essential for optimal viral replication, persistence, and pathogenicity *in vivo* (Matheson et al., 2016). Vpx was least well studied among the five accessory proteins (Vif, Vpx, Vpr, Vpu, and Nef) for its functional role and the underlying molecular basis until the recent identification of cellular anti-viral factor SAMHD1 as a target for Vpx (Hrecka et al., 2011; Laguette et al., 2011). Extensive studies since then have generated a series of critical

findings to understand the interaction of Vpx and SAMHD1 (Goldstone et al., 2011; Hrecka et al., 2011; Laguette et al., 2011, 2012; Baldauf et al., 2012; Descours et al., 2012; Lahouassa et al., 2012; Lim et al., 2012; St Gelais et al., 2012; Lenzi et al., 2014, 2015): (1) all Vpx and some Vpr proteins derived from various HIV/SIVs target SAMHD1 for proteasomal degradation; (2) SAMHD1 reduces cellular deoxynucleoside triphosphate (dNTP) pools to a level similar to that observed in non-dividing myeloid and resting T cells; (3) HIV-1 reverse transcriptase (RT) shows a high binding affinity to dNTPs relative to those from other lentiviruses. Of note here, HIV-1 can replicate in macrophages to some extent, and lacks SAMHD1-degrading activity. Thus, HIV-1 appears to be unique among primate lentiviruses to act against SAMHD1.

On the basis of the experimental results summarized above, it would hold true to hypothesize that some specific viruses have adapted themselves to better fit the physiological environments in non-dividing cells. Therefore, we here have performed phylogenetic analyses using distinct Vpx/Vpr and RT proteins from viruses with/without anti-SAMHD1 activity. To this end, we phylogenetically examined the target viral proteins (Vpx/Vpr and RT with no sequence ambiguity) derived from three types of diverse SIVs based on their genome features relating to *vpx*, *vpr*, and *vpu* genes (Fujita et al., 2010; Sakai et al., 2016) as follows. (1) "Prototype viruses" carrying *vpr* gene analyzed here were SIVagm (isolated from the African green monkey), SIVmnd-1 (mandrill), SIVlst (l'Hoest's monkey), SIVsun (sun-tailed monkey), SIVsyk (Sykes' monkey), SIVdeb (DeBrazza's monkey), SIVtal (talapoin monkey), SIVasc (red-tailed guenon), SIVcol (colobus monkey), SIVwrc (western red colobus), and SIVolc (olive colobus). (2) "HIV-1 type viruses" carrying *vpr* and *vpu* genes analyzed here were, SIVcpz (chimpanzee), SIVgor (gorilla), SIVgsn (greater spot-nosed monkey), SIVmon (mona monkey), SIVmus (mustached monkey), and SIVden (Dent's monkey). (3) "HIV-2 type viruses" carrying *vpx* and *vpr* genes analyzed here were SIVsmm (sooty mangabey monkey), SIVmac (macaque monkey), SIVmne (pig-tailed macaque), SIVstm (stump-tailed macaque), SIV-rcm (red-capped mangabey), SIVmnd-2 (mandrill), and SIVdrl (drill monkey).

PHYLOGENY OF VPX/VPR PROTEINS DERIVED FROM VIRUSES WITH/WITHOUT SAMHD1-DEGRADING ACTIVITY

Early studies have shown that Vpx is essential for HIV-2 and SIVmac replication in primary non-dividing cells (Fujita et al., 2010; Schaller et al., 2014). Subsequent studies have revealed that Vpx enhances viral DNA synthesis by inducing proteasomal degradation of an anti-viral factor in those cells (Fujita et al., 2010; Schaller et al., 2014). It is now established that the anti-viral factor SAMHD1 abundantly present in the non-dividing cells is degraded by Vpx from HIV-2 type viruses and Vpr from some prototype viruses (Laguette et al., 2012; Lim et al., 2012). There are no Vpr proteins with SAMHD1-degrading activity derived from HIV-1 and HIV-2 type viruses except for Vpr from SIVmus (HIV-1 type) (Laguette et al., 2012; Lim et al., 2012).

In order to easily see the genetic background for the results described above (Laguette et al., 2012; Lim et al., 2012), we also inferred a bootstrap phylogenetic tree of Vpx/Vpr proteins from diverse HIV/SIVs by the neighbor-joining method as previously described (Sakai et al., 2016), and the proteins with/without SAMHD1-degrading activity were highlighted by blue and red letters, respectively. Viruses without Vpx (prototype and HIV-1 type viruses) were analyzed for their Vpr proteins, and viruses with both Vpx and Vpr (HIV-2 type viruses) were examined for their Vpx proteins. As shown in **Figure 1**, Vpx/Vpr proteins from viruses with SAMHD1-degrading activity (viral groups: HIV-2, SIVsmm/mac/mne/stm, SIVrcm/mnd-2/drl, SIVagm, SIVmus/gsn/den/mon, and SIVdeb/syk/tal/asc) clearly formed different clusters from those by Vpr proteins from viruses without SAMHD1-degrading activity (viral groups: HIV-1, SIVcpz/gor, SIVmnd-1/lst/sun, and SIVolc/wrc/col). No virus strains without SAMHD1-degrading activity were found in the clusters with the activity and the opposite was true. These results suggested that HIV/SIVs with/without anti-SAMHD1 activity diverged at some time point in the past.

PHYLOGENY OF RT PROTEINS DERIVED FROM VIRUSES WITH/WITHOUT SAMHD1-DEGRADING ACTIVITY

Human immunodeficiency virus type 1 does not encode any anti-SAMHD1 proteins as described above, which appears to be disadvantageous to the virus for efficient replication in target cells, especially in non-dividing myeloid and resting T cells. Nonetheless, HIV-1 grows well in humans and is markedly pathogenic for humans. While the mechanism(s) underlying for this viral property remains elusive, one could postulate, as a likely possibility, that the high growth ability of HIV-1 is attributable to its relatively high RT activity (Schaller et al., 2014). Indeed, it has been reported that HIV-1 RT is more active in synthesizing viral DNA than the other viral RT proteins derived from viruses with SAMHD1-degrading ability (Lenzi et al., 2014, 2015).

We therefore constructed a phylogenetic tree of various RT proteins from viruses with/without anti-SAMHD1 activity in order to predict their evolutional positions as described above (Sakai et al., 2016), and the result was shown in **Figure 2**. RT proteins from viruses without anti-SAMHD1 activity (viral groups: HIV-1, SIVcpz/gor, SIVmnd-1/lst/sun, and SIVolc/wrc/col) and those from viruses with the activity (viral groups: SIVrcm/mnd-2/drl, SIVagm, SIVmus/gsn/den/mon, SIVdeb/syk/tal/asc, HIV-2, and SIVsmm/mac/mne/stm) separately formed clusters as observed in the Vpx/Vpr tree (**Figure 2**). Importantly, viruses with a high RT activity (Lenzi et al., 2014, 2015) were confined to the HIV-1 cluster without SAMHD1-degrading ability. In contrast, viruses with a low RT activity (Lenzi et al., 2014, 2015) were found in various

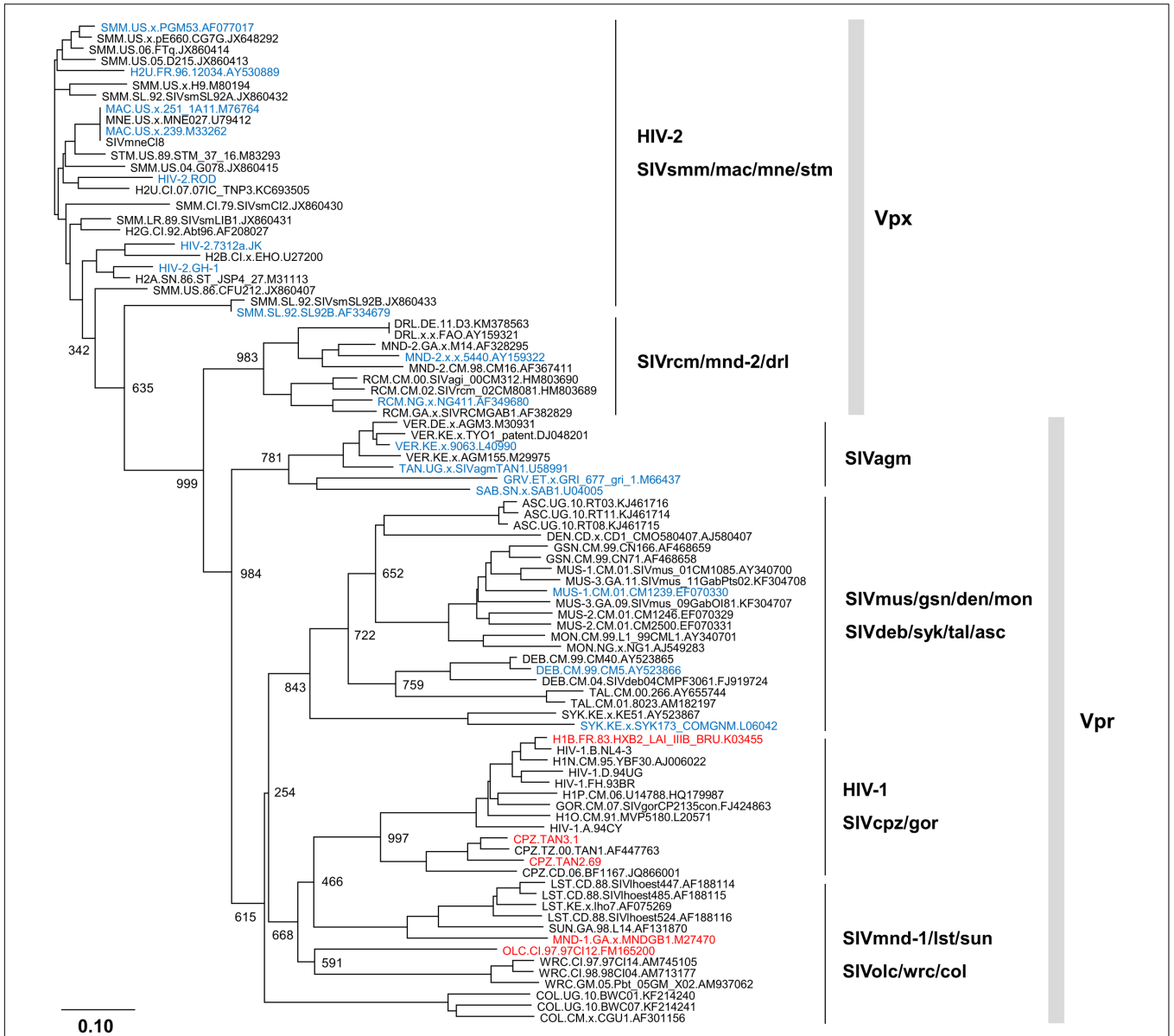


FIGURE 1 | Phylogeny of Vpx/Vpr proteins derived from various human immunodeficiency viruses/simian immunodeficiency viruses (HIV/SIVs). Phylogenetic tree based on diverse Vpx/Vpr proteins was constructed as previously described (Sakai et al., 2016), and is shown together with SAMHD1 (sterile alpha motif and HD-domain containing protein 1)-degrading activity. Viral clones with/without SAMHD1-degrading activity (Laguette et al., 2012; Lim et al., 2012) are indicated by blue and red letters, respectively. Note that Vpr proteins from HIV-2 type viruses (HIV-2, SIVsmm/mac/mne/stm, and SIVrcm/mnd-2/drl) are excluded from this tree for simplicity and clarity. Branches were calculated from 1000 bootstrap replicates, and the bootstrap values are labeled on the major branches. Scale bar represents the genetic distance. For virus designations, origins, and types, see the text (Introduction). Amino acid sequences were obtained from the HIV sequence database at Los Alamos National Laboratory (http://www.hiv.lanl.gov) or from the GenBank (http://www.ncbi.nlm.nih.gov).

virus clusters with SAMHD1-degrading ability (SIVagm, HIV-2, SIVmac, and SIVmne). While not included in this phylogenetic tree due to the sequence unavailability to us, two more HIV-1 strains (92RW and 93IN) and two other HIV-2/SIVmne strains (Rod10 and 170) were reported to have RT proteins with a high and low enzymatic activity, respectively (Lenzi et al., 2014). Together with the result in **Figure 1**, our phylogenetic tree here is consistent with a hypothesis that the ability of Vpx/Vpr proteins from various HIV/SIVs to degrade SAMHD1 and the different enzymatic activity of RT proteins from various HIV/SIVs are intimately linked.

CONCLUDING REMARKS

Primate lentiviral diversification is frequently accompanied by functional alterations of viral proteins required for efficient viral replication in different cellular environments, which can provide

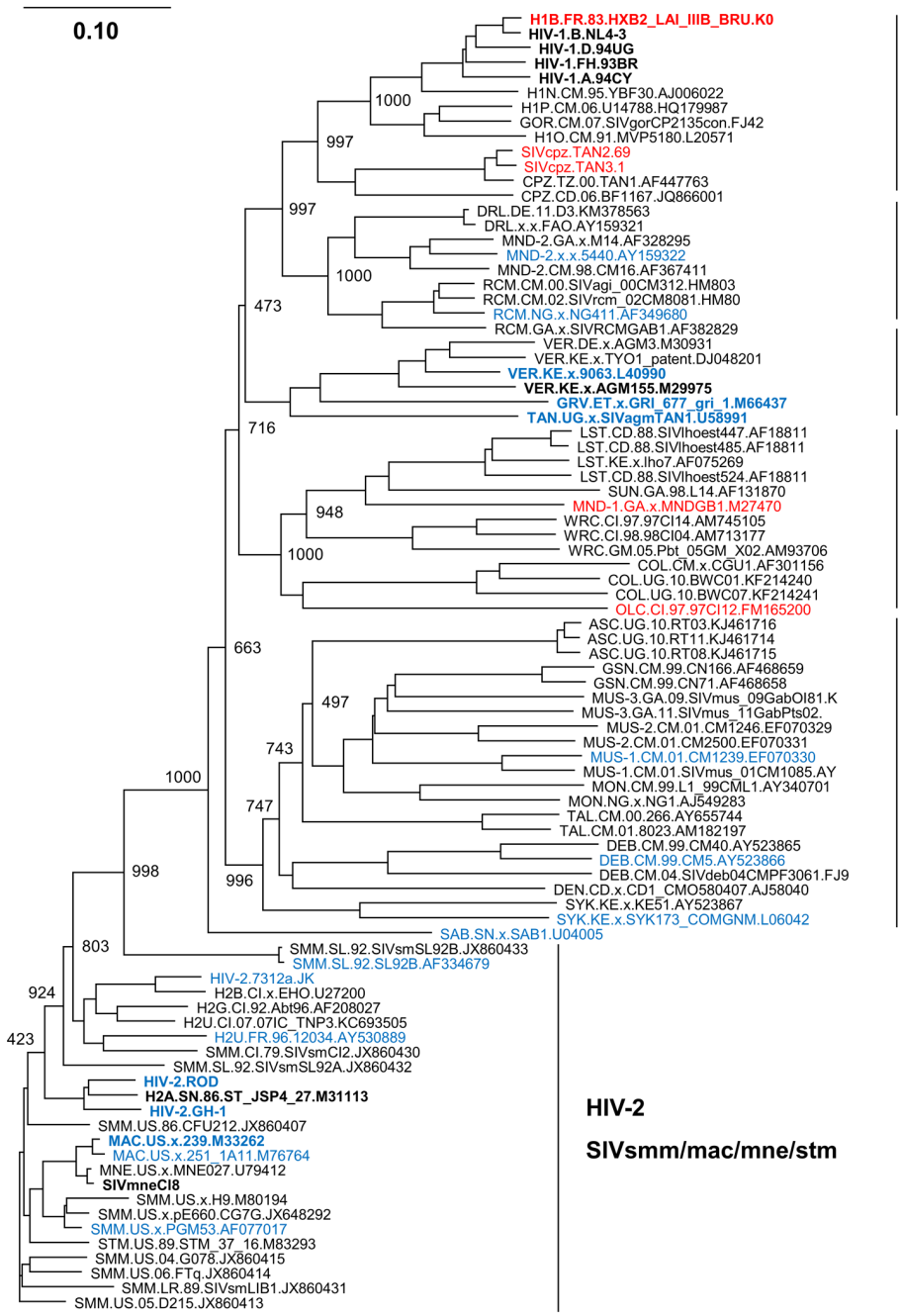


FIGURE 2 | Phylogeny of reverse transcriptase (RT) proteins derived from various HIV/SIVs. Phylogenetic tree based on diverse RT proteins was constructed as previously described (Sakai et al., 2016), and is shown together with SAMHD1-degrading and RT activities. Viral clones with/without SAMHD1-degrading activity (Laguette et al., 2012; Lim et al., 2012) are indicated by blue and red letters, respectively. In addition, viral clones experimentally tested for RT activity (Lenzi et al., 2014, 2015) are indicated by boldface type letters. Note that RT proteins from the HIV-1 group show a higher enzymatic activity relative to those from the other viral groups examined (Lenzi et al., 2014, 2015). Branches were calculated from 1000 bootstrap replicates, and the bootstrap values are labeled on the major branches. Scale bar represents the genetic distance. For virus designations, origins, and types, see the text (Introduction). Amino acid sequences were obtained from the HIV sequence database at Los Alamos National Laboratory (http://www.hiv.lanl.gov) or from the GenBank (http://www.ncbi.nlm.nih.gov).

virologically essential informations to understand the inter-/intra-species transmission and adaptation processes. In this report, we have described a potential link between the SAMHD1-degrading activity of Vpx/Vpr proteins and the enzymatic activity of RT proteins. Our phylogenetic analyses (**Figures 1** and **2**) are consistent with a scenario that HIV-1 without anti-SAMHD1 activity may already have or have acquired RT with enhanced enzymatic activity. However, because viral strains experimentally

examined for their anti-SAMHD1 and RT activities through biochemical and molecular biological approaches are limited so far, it would be too early to draw a clear conclusion on this issue. Extensive experimental studies are required to obtain decisive results.

AUTHOR CONTRIBUTIONS

YS, AA, and MN designed the research. YS performed the phylogenetic analysis. YS, ND, YM, AA, and MN analyzed and discussed the results. YS, AA, and MN wrote the manuscript.

FUNDING

This study is supported in part by a research grant from the Japan Society for the Promotion of Science (JSPS KAKENHI: Grant Number JP26293104).

ACKNOWLEDGMENT

We thank Ms. Kazuko Yoshida (Department of Microbiology, Tokushima University Graduate School of Medical Science, Tokushima, Japan) for her editorial assistance.

REFERENCES

Baldauf, H. M., Pan, X., Erikson, E., Schmidt, S., Daddacha, W., Burggraf, M., et al. (2012). SAMHD1 restricts HIV-1 infection in resting CD4+ T cells. *Nat. Med.* 18, 1682–1687. doi: 10.1038/nm.2964

Descours, B., Cribier, A., Chable-Bessia, C., Ayinde, D., Rice, G., Crow, Y., et al. (2012). SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells. *Retrovirology* 9:87. doi: 10.1186/1742-4690-9-87

Fujita, M., Otsuka, M., Nomaguchi, M., and Adachi, A. (2010). Multifaceted activity of HIV Vpr/Vpx proteins: the current view of their virological functions. *Rev. Med. Virol.* 20, 68–76. doi: 10.1002/rmv.636

Goldstone, D., Ennis-Adeniran, V., Hedden, J., Groom, H., Rice, G., Christodoulou, E., et al. (2011). HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* 480, 379–382. doi: 10.1038/nature10623

Hrecka, K., Hao, C., Gierszewska, M., Swanson, S. K., Kesik-Brodacka, M., Srivastava, S., et al. (2011). Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474, 658–661. doi: 10.1038/nature10195

Laguette, N., Rahm, N., Sobhian, B., Chable-Bessia, C., Münch, J., Snoeck, J., et al. (2012). Evolutionary and functional analyses of the interaction between the myeloid restriction factor SAMHD1 and the lentiviral Vpx protein. *Cell Host Microbe* 11, 205–217. doi: 10.1016/j.chom.2012.01.007

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Ségéral, E., et al. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474, 654–657. doi: 10.1038/nature10117

Lahouassa, H., Daddacha, W., Hofmann, H., Ayinde, D., Logue, E., Dragin, L., et al. (2012). SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* 13, 223–228. doi: 10.1038/ni.2236

Lenzi, G., Domaoal, R., Kim, D.-H., Schinazi, R., and Kim, B. (2015). Mechanistic and kinetic differences between reverse transcriptases of Vpx coding and non-coding lentiviruses. *J. Biol. Chem.* 290, 30078–30086. doi: 10.1074/jbc.M115.691576

Lenzi, G. M., Domaoal, R. A., Kim, D.-H., Schinazi, R. F., and Kim, B. (2014). Kinetic variations between reverse transcriptases of viral protein X coding and noncoding lentiviruses. *Retrovirology* 11:111. doi: 10.1186/s12977-014-0111-y

Lim, E. S., Fregoso, O. I., McCoy, C. O., Matsen, F. A., Malik, H. S., and Emerman, M. (2012). The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. *Cell Host Microbe* 11, 194–204. doi: 10.1016/j.chom.2012.01.004

Matheson, N. J., Greenwood, E. J. D., and Lehner, P. J. (2016). Manipulation of immunometabolism by HIV – accessories to the crime? *Curr. Opin. Virol.* 19, 65–70. doi: 10.1016/j.coviro.2016.06.014

Sakai, Y., Miyake, A., Doi, N., Sasada, H., Miyazaki, Y., Adachi, A., et al. (2016). Expression profiles of Vpx/Vpr proteins are co-related with the primate lentiviral lineage. *Front. Microbiol.* 7:1211. doi: 10.3389/fmicb.2016.01211

Schaller, T., Bauby, H., Hué, S., Malim, M. H., and Goujon, C. (2014). New insights into an X-traordinary viral protein. *Front. Microbiol.* 5:126. doi: 10.3389/fmicb.2014.00126

St Gelais, C., De Silva, S., Amie, S. M., Coleman, C. M., Hoy, H., Hollenbaugh, J. A., et al. (2012). SAMHD1 restricts HIV-1 infection in dendritic cells (DCs) by dNTP depletion, but its expression in DCs and primary CD4+ T-lymphocytes cannot be upregulated by interferons. *Retrovirology* 9:105. doi: 10.1186/1742-4690-9-105

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AK and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Sakai, Doi, Miyazaki, Adachi and Nomaguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 June 2017
doi: 10.3389/fmicb.2017.01082

Comparison of Biochemical Properties of HIV-1 and HIV-2 Capsid Proteins

Yasuyuki Miyazaki[1], Ariko Miyake[2], Noya Doi[3], Takaaki Koma[3], Tsuneo Uchiyama[3], Akio Adachi[3] and Masako Nomaguchi[3]**

[1] Department of Microbiology and Cell Biology, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan, [2] Laboratory of Molecular Immunology and Infectious Disease, Joint Faculty of Veterinary Medicine, Yamaguchi University, Yamaguchi, Japan, [3] Department of Microbiology, Tokushima University Graduate School of Medical Sciences, Tokushima, Japan

Timely disassembly of viral core composed of self-assembled capsid (CA) in infected host cells is crucial for retroviral replication. Extensive *in vitro* studies to date on the self-assembly/disassembly mechanism of human immunodeficiency virus type 1 (HIV-1) CA have revealed its core structure and amino acid residues essential for CA–CA intermolecular interaction. However, little is known about *in vitro* properties of HIV-2 CA. In this study, we comparatively analyzed the polymerization properties of bacterially expressed HIV-1 and HIV-2 CA proteins. Interestingly, a much higher concentration of NaCl was required for HIV-2 CA to self-assemble than that for HIV-1 CA, but once the polymerization started, the reaction proceeded more rapidly than that observed for HIV-1 CA. Analysis of a chimeric protein revealed that N-terminal domain (NTD) is responsible for this unique property of HIV-2 CA. To further study the molecular basis for different *in vitro* properties of HIV-1 and HIV-2 CA proteins, we determined thermal stabilities of HIV-1 and HIV-2 CA NTD proteins at several NaCl concentrations by fluorescent-based thermal shift assays. Experimental data obtained showed that HIV-2 CA NTD was structurally more stable than HIV-1 CA NTD. Taken together, our results imply that distinct *in vitro* polymerization abilities of the two CA proteins are related to their structural instability/stability, which is one of the decisive factors for viral replication potential. In addition, our assay system described here may be potentially useful for searching for anti-CA antivirals against HIV-1 and HIV-2.

Keywords: HIV-1, HIV-2, Gag-CA, NTD, CA-polymerization, CA-stability

OPEN ACCESS

Edited by:
Akihide Ryo,
Yokohama City University, Japan

Reviewed by:
Subrata H. Mishra,
Johns Hopkins School of Medicine, United States
Jamil S. Saad,
University of Alabama at Birmingham, United States

****Correspondence:***
Akio Adachi
adachi@tokushima-u.ac.jp
Masako Nomaguchi
nomaguchi@tokushima-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *27 March 2017*
Accepted: *29 May 2017*
Published: *13 June 2017*

Citation:
Miyazaki Y, Miyake A, Doi N, Koma T, Uchiyama T, Adachi A and Nomaguchi M (2017) Comparison of Biochemical Properties of HIV-1 and HIV-2 Capsid Proteins. Front. Microbiol. 8:1082. doi: 10.3389/fmicb.2017.01082

INTRODUCTION

Highly ordered core structure of human immunodeficiency virus type 1 (HIV-1) consisting of multimeric capsid (CA) proteins is essential for modulating the complex virus replication (Freed and Martin, 2013; Goff, 2013). While unusual stabilization by mutations in CA abrogates viral infectivity through incomplete reverse transcription of viral genome (Forshey et al., 2002), rhesus α-isoform of tripartite motif-containing protein 5 (TRIM5α) eliminates viral infectivity by abnormally promoting disassembly of CA proteins (Forshey et al., 2005; Sebastian and Luban, 2005; Stremlau et al., 2006). A variety of host proteins have been reported to regulate CA disassembly: cyclophilin A (CypA) (Braaten et al., 1996a,b; Gamble et al., 1996), PDZ domain-containing protein 8 (PDZD8) (Henning et al., 2010; Guth and Sodroski, 2014), cleavage and polyadenylation

specificity factor (CPSF) (Lee et al., 2010; Price et al., 2012), and myxovirus resistance protein 2 (MX2) (Goujon et al., 2013; Kane et al., 2013; Liu et al., 2013). Thus, proper disassembly of the CA-core structure in the viral replication cycle in concert with cellular proteins is critical for HIV-1 infectivity. Given that the process of HIV-1 CA core dissociation in infected cells is intricately mediated by numerous viral and cellular factors, *in vitro* model systems that mimic the *in vivo* situation to a certain extent are required to gain definite insights into molecular events in HIV-1 core formation/deformation. In fact, various *in vitro* systems have been developed to study the physicochemical aspects of HIV-1 CA–CA interaction (Ehrlich et al., 1992; Campbell and Vogt, 1995; Gross et al., 1997; von Schwedler et al., 1998; Gross et al., 2000; Ehrlich et al., 2001; Lanman et al., 2002; Mateu, 2002; Morikawa et al., 2004; Alfadhli et al., 2005; del Alamo et al., 2005; Lidon-Moya et al., 2005; Chen and Tycko, 2010). Although the above systems are influenced by numerous factors, such as ion strength, temperature, pH, and crowding agents, a high concentration of NaCl has been generally and frequently used to initiate the CA-assembly *in vitro*.

HIV CA consists of two globular domains [N-terminal domain (NTD) and C-terminal domain (CTD)], and a linker domain connecting these two domains (**Figure 1A**). While HIV-1 CA NTD has an N-terminal β-hairpin and seven α-helices in the downstream region (Gitti et al., 1996; Momany et al., 1996), its CA CTD has four α-helices (Gamble et al., 1997; Du et al., 2011). The former primarily forms a hexameric structure (also forms a pentameric structure), and the latter interacts with the NTD of adjacent CA molecules (Pornillos et al., 2009, 2011). CTD also forms a CTD-CTD dimer (Ganser-Pornillos et al., 2007; Pornillos

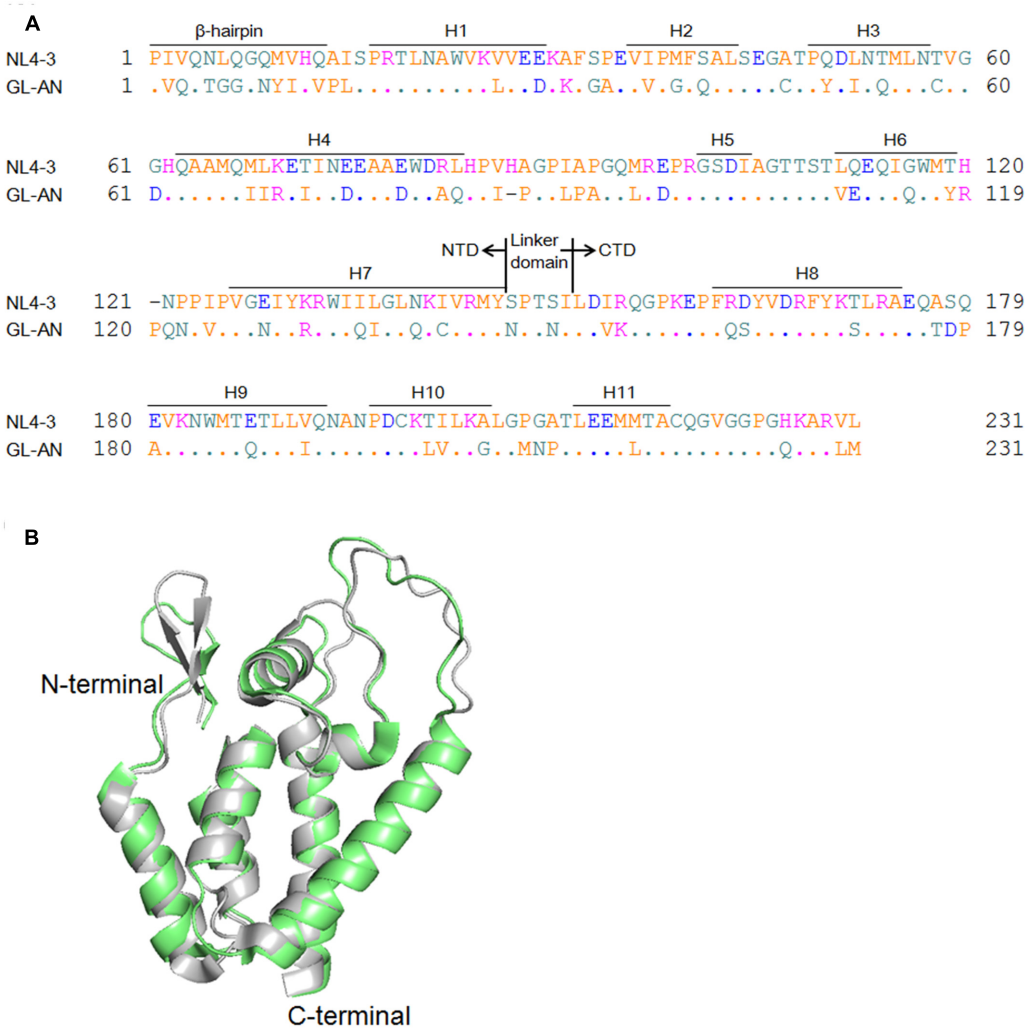


FIGURE 1 | Structural comparison of NL4-3 and GL-AN Gag-CA proteins. **(A)** Alignment of NL4-3 and GL-AN Gag-CA sequences. CA amino acid sequences of HIV-1 NL4-3 (GenBank accession number: AF324493) and HIV-2 GL-AN (GenBank accession number: M30895) are aligned. The N-terminal domain (NTD), linker domain, C-terminal domain (CTD), β-hairpin, and helices 1 to 11 (H1 to H11) are shown based on previous studies (Gamble et al., 1996; von Schwedler et al., 2003; Robinson et al., 2014; Gres et al., 2015). **(B)** Superimposition of the NTD structures. Superposed structures of HIV-1 NTD (green, PDB code: 3H4E) and HIV-2 NTD (gray, PDB code: 2WLV) were depicted by PyMOL ver 1.8.

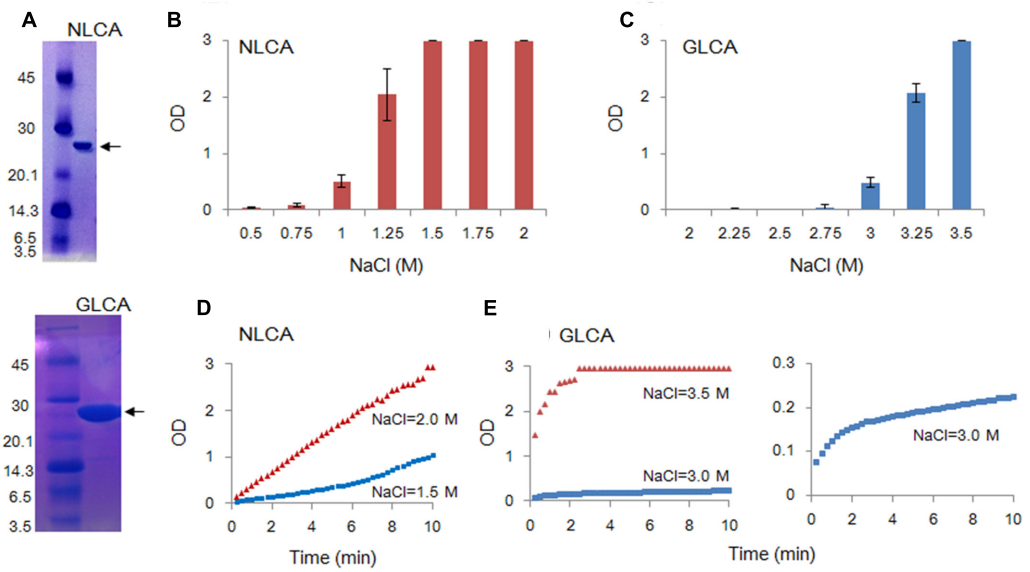


FIGURE 2 | *In vitro* polymerization features of NL4-3 and GL-AN CA proteins. Polymerization reaction was carried out and monitored by OD at 350 nm as described in MATERIALS AND METHODS. **(A)** SDS-PAGE profile. The purity of CA proteins was checked by SDS-PAGE gel stained with Coomassie Brilliant Blue. Size markers in kDa (on the left) and the CA proteins (arrows) are indicted. NLCA, NL4-3 CA; GLCA, GL-AN CA. **(B–E)** Self-polymerization of CA proteins. Purified CA proteins were added to the buffer containing various concentrations of NaCl, and incubated for 4 h **(B–C)** or 10 min (**D–E)** at room temperature. Polymerized products and polymerization kinetics at various concentrations of NaCl were determined by OD at 350 nm for NLCA **(B,D)** and GLCA **(C,E)** as shown.

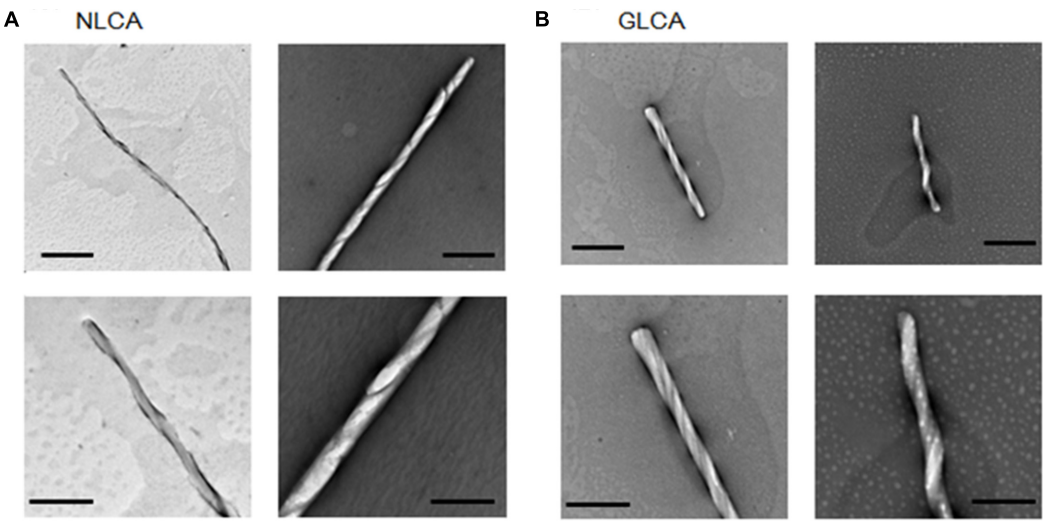


FIGURE 3 | Tubular structures of *in vitro* assembled NL4-3 and GL-AN CA proteins. CA proteins were fully polymerized *in vitro*, fixed by glutaraldehyde, and visualized via TEM as previously described (Sakaguchi et al., 2002; Piroozmand et al., 2006). **(A,B)** represent micrographs of HIV-1 NL4-3 CA (NLCA) and HIV-2 GL-AN CA (GLCA), respectively. Scale bars: 1 μm in upper panels; 200 nm in lower panels.

et al., 2009). Based on cryo-electron microscopy studies, it has been proposed that HIV-1 cores in mature virions have a mixture of approximately 250 hexameric CA proteins and 12 pentameric CA proteins at upper and lower ends of the conical core (Ganser et al., 1999; Zhao et al., 2013). *In vitro* studies have also shown that HIV-1 CA is assembled to form core-like structure made up of hexameric CA proteins, a structure similar to the core in native virions (Ganser et al., 1999; Byeon et al., 2009; Zhao et al., 2013). Of note, this self-assembly process of HIV-1 CA (monomers, hexamers, and final core-like products consisting of hexamers) can be induced by high ionic strength, and readily monitored by simple turbidity assays *in vitro* (Ehrlich et al., 1992; Li et al., 2000; Ganser-Pornillos et al., 2004; Barklis et al., 2009). Moreover, as described above, CA disassembly process regulated by CA inhibitors/host proteins, such as TRIM5α, CypA, and PDZD8, can be experimentally analyzed *in vitro* as well as *in vivo*

(Grattinger et al., 1999; Ternois et al., 2005; Black and Aiken, 2010; Guth and Sodroski, 2014).

Amino acid sequences of HIV-1/HIV-2 CA proteins are significantly related to each other (**Figure 1A**), and more strikingly, their NTD 3-D structures are highly similar (**Figure 1B**). However, although HIV-1 and HIV-2 exhibit distinct biological properties associated with their CA proteins (Freed and Martin, 2013), to the best of our knowledge, *in vitro* properties of HIV-2 CA have been very poorly studied so far. HIV-2 is a medically and socially important retrovirus in addition to HIV-1, and is important for basic virology as well. In this study, we comparatively analyzed the *in vitro* polymerization properties of HIV-1/HIV-2 CA proteins, and also their thermal stability. We found that HIV-1 and HIV-2 CA proteins are remarkably different from each other in these characteristics, and demonstrated that the observed difference is attributable to the NTD of CA proteins. Our results here suggest that the structural instability/stability of CA NTD influences distinct biological properties of HIV-1 and HIV-2.

MATERIALS AND METHODS

Plasmids

Sequences encoding a full-length CA of HIV-1 NL4-3 (Pro1-Leu231 in **Figure 1A**) and its NTD (Pro1-Tyr145 in **Figure 1A**) were PCR-amplified and cloned into pET21 (EMD chemicals, Inc.) using *Nde* I and *Xho* I sites to generate NLCA and NLNTD, respectively. Sequences encoding a full-length CA of HIV-2 GL-AN (Pro1-Met231 in **Figure 1A**) and its NTD (Pro1-Tyr145 in **Figure 1A**) were PCR-amplified and cloned into pET21 as above to generate GLCA and GLNTD, respectively. All mutant clones analyzed in this study, designated NL/GL, GL32NLCA, and GLmtCA, were generated by overlapping PCR. Infectious molecular clones designated NL4-3 and GL-AN have been previously described (Adachi et al., 1986; Shibata et al., 1990; Kawamura et al., 1994). NL4-3 and GL-AN are prototype full-length clones of HIV-1 and HIV-2, respectively, and frequently and widely used for various HIV studies as representative clones.

Expression and Purification of CA Proteins

Recombinant CA proteins (tagged with poly-histidine at the C-terminus) were expressed in *E. coli* strain BL21 (DE3) RIL (Agilent technologies) and purified as previously described (Li et al., 2000). Expression of wild-type and mutant CA proteins were induced with 0.1 mM isopropyl β-D-thiogalactopyranoside. Recombinant proteins were then purified by immobilized metal affinity chromatography (TALON, Clontech Laboratories Inc.), and their purities were checked and confirmed by SDS-poly-acrylamide gel electrophoresis (SDS-PAGE) for subsequent experimental uses (**Figure 2A**).

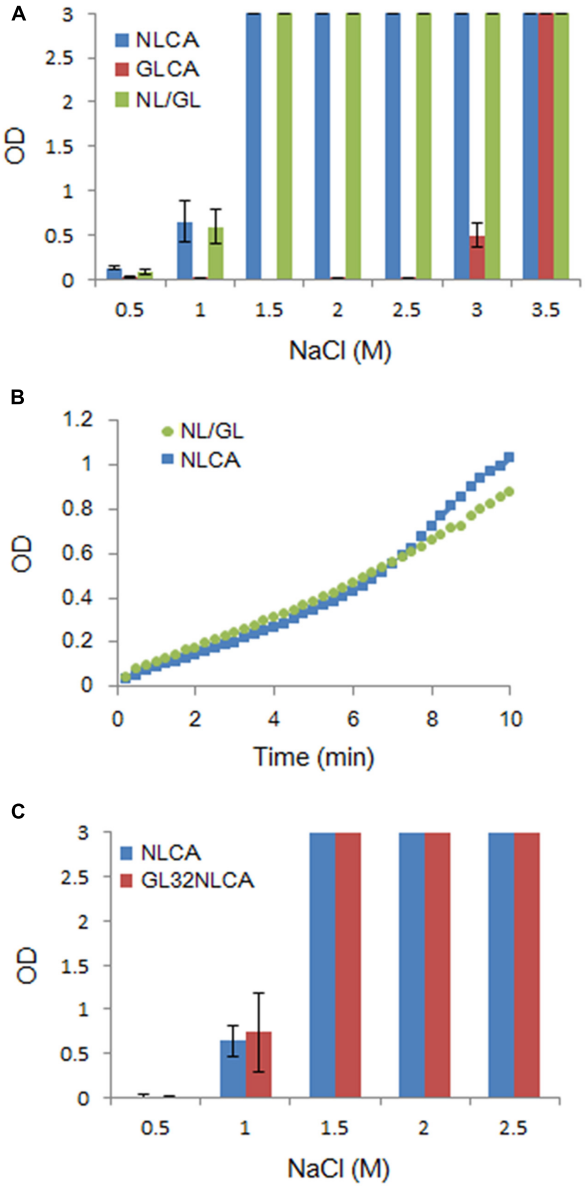


FIGURE 4 | Comparative analysis of several CA proteins for their *in vitro* polymerization properties. CA-polymerization was performed *in vitro* and monitored by OD at 350 nm as described in MATERIALS AND METHODS. **(A)** Polymerization of NL4-3 CA (NLCA), GL-AN CA (GLCA), and NL/GL CA (NL/GL) proteins for 4 h at various NaCl concentrations. The chimeric clone NL/GL has the sequence encoding the NTD of NL4-3 CA and the linker domain/CTD of GL-AN CA (**Figure 1A**). **(B)** Polymerization kinetics of NL and NL/GL CA proteins (1.5 M NaCl). **(C)** Polymerization of NL and GL32NL CA proteins for 4 h at various NaCl concentrations. GL32NL is a chimeric NLCA-derivative clone which has the sequence encoding the very N-terminal region of GL-AN CA (Pro1-Phe32 in **Figure 1A**).

In Vitro Assay for HIV CA Assembly

Assays for CA assembly were performed similarly as described previously (Ehrlich et al., 1992; Li et al., 2000; Ganser-Pornillos et al., 2004; Barklis et al., 2009). Three independent polymerization experiments were carried out, and mean

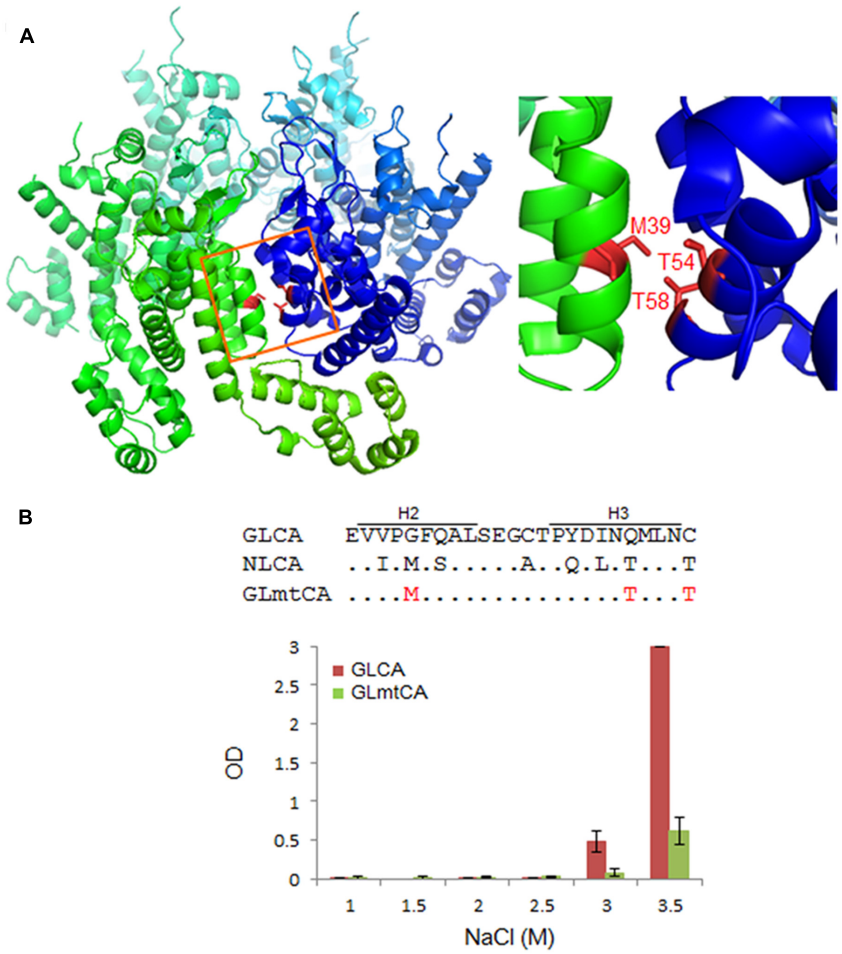


FIGURE 5 | Comparison of *in vitro* polymerization activity of GL-AN and mutant CA proteins. **(A)** CA–CA interacting surface of HIV-1. Two CA molecules (greenish and bluish) are shown. Three amino acids (M39, T54, and T58 in the H2–H3 region) critical for the interaction are highlighted as shown (see Pornillos et al., 2009, 2011 for details). The structure of HIV-1 CA (PDB code: 3H4E) was depicted by PyMOL ver 1.8. **(B)** Polymerization of parental and mutant CA proteins. CA-polymerization was performed *in vitro* for 4 h at various NaCl concentrations, and monitored by OD at 350 nm as described in the Section "Materials and Methods". The mutant clone GLmtCA has three amino acid substitutions relative to GLCA as indicated.

values ± standard deviations are shown where indicated. CA polymerization reactions were performed in 50 mM Tris-HCl (pH 8.0) at a final concentration of 50 μM recombinant CA proteins. Reactions were carried out at different NaCl concentrations for the indicated time in each experiment. Polymerized products were monitored by optical density (OD) at 350 nm using a spectrophotometer (DU730, Beckman–Coulter). OD_{350} has been used to most sensitively measure the ordered protein aggregates, not the protein itself. The upper OD detection limit of DU730 is 3.0.

Transmission Electron Microscopy (TEM)

Negative staining and electron microscopy were performed similarly as described before (Sakaguchi et al., 2002; Piroozmand et al., 2006). NL4-3 and GL-AN CA proteins adjusted to 100 μM were *in vitro* polymerized in the presence of 2 M NaCl (for NL4-3) or 3.5 M NaCl (for GL-AN) for 30 min at room temperature. Assembled CA proteins were fixed by 0.2% glutaraldehyde, and were placed on formvar-carbon-coated nickel grids, stained with 4% uranyl acetate, and examined by a transmission electron microscope (Hitachi H-7650).

Fluorescence-based Thermal Shift Assay

Fluorescence-based thermal shift assays by differential scanning fluorimetry (DSF) were carried out as described previously (Niesen et al., 2007; Fedorov et al., 2012). Three independent DSF experiments were conducted with highly similar results. Fifty μM of CA NTD proteins were prepared in 50 mM Tris-HCl (pH 8.0), 250–2000 mM NaCl and 1 mM 2-mercaptethanol containing SYPRO orange (Invitrogen) to quantify thermal stability. Temperature gradient was set in the range of 25°C to 95°C with 1% ramp rate using 7500 real-time PCR system (Applied Biosystems). Melting temperature (Tm) of CA was calculated by SYPRO orange fluorescence curves using 7500 software ver. 2.03.

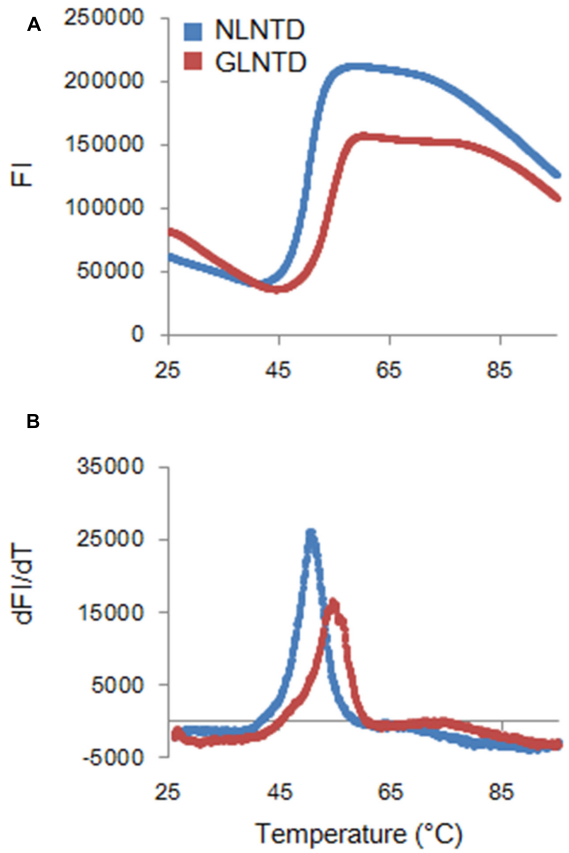


FIGURE 6 | Thermal stability of CA NTD proteins derived from HIV-1 NL4-3, and HIV-2 GL-AN. The thermal stability of NLCA and GLCA NTD proteins in the presence of 250 mM NaCl was determined by DSF as described in the Section "Materials and Methods". SYPRO orange fluorescence intensity (FI) at varying temperatures (upper panel) and derivative melt curves calculated by differences in FI at each temperature (lower panel) are shown. Peak temperatures in the curves (dFI/dT) were considered as Tm.

RESULTS

Polymerization Properties of HIV-1 and HIV-2 CA Proteins

Capsid protein is a major component of retroviral particles, and commonly plays an essential role for virus replication (Goff, 2013). Therefore, amino acid sequences in CA proteins and their structural features are expected to be conserved among viruses, especially those belonging to the same viral genus. Indeed, when HIV-1 and HIV-2 CA proteins are compared (**Figure 1**), their amino acid identities are around 70% (67% for NTD and 73% for CTD), and structural outlines of NTD proteins as revealed by nuclear magnetic resonance are highly similar (Price et al., 2009). Nevertheless, we were interested in ascertaining if there is some biochemical/biophysical difference(s) between the two closely related CA molecules from HIV-1 and HIV-2 that may potentially affect their biological properties.

To comparatively determine the physicochemical characteristics of HIV-1 and HIV-2 CA proteins, we employed the *in vitro* assembly system by NaCl (Ehrlich et al., 1992; Li et al., 2000; Ganser-Pornillos et al., 2004; Barklis et al., 2009) in this study. Recombinant CA proteins (derived from HIV-1 NL4-3 and HIV-2 GL-AN) expressed in bacteria and purified were used for monitoring their polymerization reactions. As shown in **Figures 2B,C**, a very prominent difference was noted between NL4-3 and GL-AN CA proteins for their NaCl-dependent polymerization. NL4-3 CA polymerized at a significant level with 1 M NaCl, and fully assembled with 1.5 M, 1.75 M, and 2 M NaCl. In a sharp contrast, no polymerized products were detected for GL-AN CA with 2.75 M and lower concentrations of NaCl. At least 3 M NaCl or higher concentrations were required for GL-AN CA to definitively detect the polymerization products. We then performed kinetic studies to determine whether there was any difference between NL4-3 and GL-AN CA proteins (**Figures 2D,E**). As demonstrated, NL4-3 CA self-assembled in a linear way over time with 1.5 M and 2.0 M NaCl, whereas GL-AN CA exhibited Boltzmann shape curves with 3 M and 3.5 M NaCl even when OD values were very low (note the GLCA right panel in **Figure 2E**). In total, GL-AN CA self-polymerized more rapidly than NL4-3 CA once successfully triggered by NaCl. Considering early reports on the TEM morphology of *in vitro* HIV-1 CA products (Ehrlich et al., 1992; Campbell and Vogt, 1995; Gross et al., 1997), we checked for the presence of assembled CA proteins with a tubular or cylinder shape in the reaction products. As shown in **Figures 3A,B**, polymers with a similar morphology were readily observed in the *in vitro* assembled products of NL4-3 and GL-AN CA proteins, as previously described.

Polymerization Properties of CA Chimeric and Mutant Proteins

Although the amino acid identities in NL4-3 and GL-AN CA proteins are considerably high (**Figure 1**), their polymerization properties *in vitro* were clearly distinct (**Figure 2**). To determine the region in CA responsible for the observed differences, we first constructed a chimeric clone between the two CA proteins and designated it as NL/GL. NL/GL contains the NTD of NL4-3 CA and the linker/CTD of GL-AN CA (**Figure 1A**). We then performed the *in vitro* polymerization assays for NL4-3 CA, GL-AN CA, and NL/GL as described above. As clearly seen in **Figure 4A**, NL/GL gave very similar results with those obtained for NL4-3 CA, but very different from those for GL-AN CA. Consistent with these results, NL/GL polymerization kinetics were highly similar to those of NL4-3 CA (**Figure 4B**). We concluded that NTD determines differences in *in vitro* polymerization properties of NL4-3 and GL-AN CA proteins.

It has been reported that the N-terminal β-hairpin structure is important for CA assembly by analyzing a chimeric protein between HIV-1 and murine leukemia virus CA proteins (Cortines et al., 2011). This suggested that the most N-terminal region containing the β-hairpin structure could be determinants for the different polymerization phenotypes between NL4-3 and GL-AN CA proteins. We therefore constructed a chimeric CA clone designated GL32NLCA to test this possibility. The N-terminal portion of NLCA (Pro1-Phe32 in **Figure 1A**) was replaced with the corresponding region of GLCA to generate GL32NLCA.

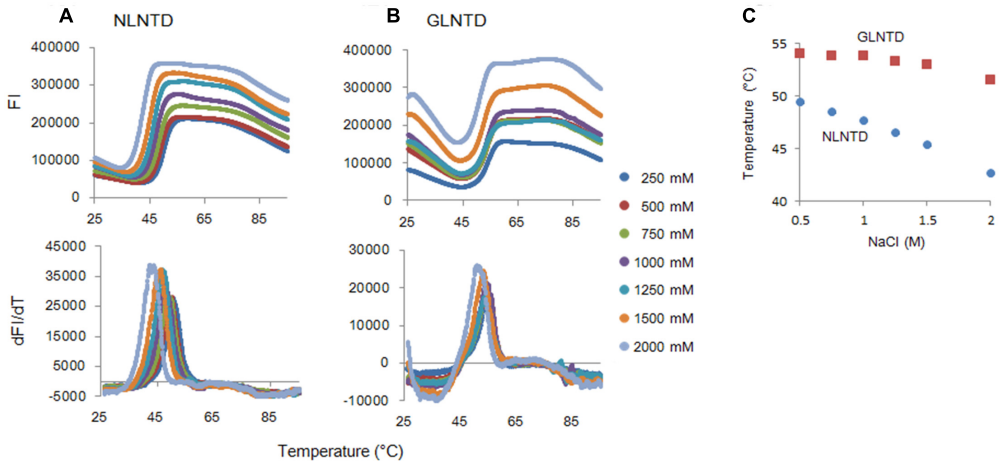


FIGURE 7 | Thermal stability of CA NTD proteins at various NaCl concentrations. The thermal stability of NLCA and GLCA NTD proteins at various NaCl concentrations was evaluated as described in the legend to **Figure 6**. FI values and melt curves (**A,B**, respectively), and Tm shifts induced by NaCl **(C)** are shown.

As shown in **Figure 4C**, when NLCA and GL32NLCA were monitored for the polymerization activity at various NaCl concentrations, no significant difference was noted. This result indicated that the region containing the N-terminal β-hairpin structure is not a determinant for the distinct polymerization properties.

Previous structural studies have shown that some amino acid residues (no. 20, 38, 39, 42, 54, and 58 in CA) located at the CA–CA interaction interface (regions of helices 1, 2, and 3) are critical for polymerization (Pornillos et al., 2009, 2011). Of these six residues, three, i.e., L20, P38, and A42, are conserved between NL4-3 and GL-AN CA proteins (**Figure 1A**). Consequently, three amino acids (no. 39, 54, and 58) are unshared, and reside in the CA–CA interacting surface (**Figure 5**). We therefore introduced three amino acid substitutions (G39M, Q54T, and C58T) into GLCA to generate GLmtCA carrying residues of the NL4-3 CA type (**Figure 5**), and examined its polymerization ability. GLmtCA was found to have much less polymerization activity than GLCA (**Figure 5B**). This result indicated that amino acids G39, Q54, and C58 are critical for self-assembly of GL-AN CA, and that the amino acid substitutions introduced cannot change its polymerization property to the NL4-3 type.

Thermal Stability of CA NTD Proteins

It has been reported that chemical chaperons, inhibitors of HIV-1 CA polymerization, raise its Tm (Lampel et al., 2013). Furthermore, previous studies have reported that higher temperature facilitates the polymerization of HIV-1 CA (Ehrlich et al., 1992; Morikawa et al., 2004; Alfadhli et al., 2005). Moreover, the interaction of HIV CA with CypA or anti-retroviral TRIM-Cyp, known to promote HIV CA dissociation (Braaten et al., 1996a,b; Gamble et al., 1996; Forshey et al., 2005; Sebastian and Luban, 2005; Stremlau et al., 2006), was shown to be an exothermic event (Yoo et al., 1997; Price et al., 2009), indicating that HIV CA is shifted to a more thermally stable state upon binding to CypA or TRIM-Cyp. These results have strongly suggested that the ability of CA to polymerize and the thermal stability of CA are mutually linked. We therefore comparatively determined the thermal stability of CA NTD proteins derived from NL4-3 and GL-AN, the determinants for distinct polymerization properties of the two CA proteins (**Figure 4**).

To determine the Tm for the two proteins, we employed fluorescence-based thermal shift assays using SYPRO orange. Proteins expose hydrophobic patches upon heating. This assay utilizes a chemical property of SYPRO orange to bind to hydrophobic patches in target proteins, and therefore, their denaturation states can be monitored by fluorescence intensities from SYPRO orange. As shown in **Figure 6**, fluorescence intensity curves of NL4-3 and GL-AN NTD proteins obtained by this assay system were quite different, and Tm values for NL4-3 NTD and GL-AN NTD were calculated to be 50.4 °C and 53.9 °C, respectively. Thus, GL-AN NTD was thermally more stable than NL4-3 NTD. We further examined the effects of NaCl, an agent to promote CA polymerization (**Figure 2**), on the thermal stability of the two NTD proteins. **Figure 7** shows the results obtained. The Tm for NL4-3 NTD fell down (7.4 °C) by increasing NaCl concentrations from 250 mM to 2000 mM (**Figures 7A,C**). On the other hand, relatively mild effects were observed for the Tm-shift of GL-AN NTD (2.1°C falling down as shown in **Figures 7B,C**). Thus, GL-AN NTD was less influenced by NaCl with respect to thermal stability than NL4-3 NTD. In other words, NL4-3 NTD was structurally destabilized to a greater extent by NaCl than GL-AN NTD. Collectively, GL-AN NTD was structurally more stable than NL4-3 NTD.

DISCUSSION

Although there have been numerous studies on the assembly of HIV-1 CA *in vitro* (Ehrlich et al., 1992, 2001; Campbell and Vogt, 1995; Gross et al., 1997, 2000; von Schwedler et al., 1998; Lanman et al., 2002; Mateu, 2002;

Morikawa et al., 2004; Alfadhli et al., 2005; del Alamo et al., 2005; Lidon-Moya et al., 2005; Chen and Tycko, 2010), almost no experimental investigations into the corresponding research field of HIV-2 have been done as of yet. In this work, we determined *in vitro* characteristics of HIV-2 CA (derived from an infectious clone, GL-AN) in parallel with HIV-1 CA (derived from an infectious clone, NL4-3). We confirmed previous findings on HIV-1 CA, and newly found that HIV-2 CA is strikingly distinct from HIV-1 CA regarding its *in vitro* properties (**Figures 2**, **4**, **6**, **7**) despite their sequence relatedness and structural similarity (**Figure 1**). We demonstrated here that much higher concentrations of NaCl are required for the polymerization of HIV-2 CA than for HIV-1 CA, but that HIV-2 CA assembly proceeds more promptly relative to HIV-1 CA after being initiated (**Figures 2**, **4**). Although a specific narrow region in CA was not identified as a determinant, NTD is clearly responsible for this property (**Figure 4**). This conclusion is quite reasonable because amino acid identities are higher in CTD than those in NTD, and would be consistent with the finding that intermolecular CTD-CTD interaction occurs first and NTD-NTD interaction occurs as the final step (Grime and Voth, 2012; Tsiang et al., 2012). NTD may function as a rate-limiting factor for CA polymerization.

We also demonstrated by fluorescence-based thermal shift assays that HIV-2 CA NTD is structurally more stable than HIV-1 CA NTD (**Figures 6**, **7**). Thus, the thermal stability of NTD proteins was inversely related with the polymerization ability of CA proteins at lower NaCl concentrations (**Figures 2**, **4**). Although the molecular basis is still unknown, the negative relationship between thermal stability and assembly property has been reported for CA proteins of HIV-1 and its CA mutant (Ganser-Pornillos et al., 2004; Cortines et al., 2015), and for HIV-1 CA treated with chemical chaperons (Lampel et al., 2013). Furthermore, HIV-1 CA polymerization was facilitated by heat destabilization (Ehrlich et al., 1992; Morikawa et al., 2004; Alfadhli et al., 2005; Barklis et al., 2009). Taken together, it is not unreasonable to assume that the relatively high thermal stability of CA is associated with the relatively poor assembly property (i.e., high disassembly property) of CA. However, whether the different sequence and/or structural property of HIV-1/HIV-2 CA proteins needed to bind to cellular restriction or regulation factors, such as TRIM5α or CypA, are linked with our results described here is unclear at present. The molecular basis underlying the distinct *in vitro* features of HIV-1/HIV-2 CA proteins remains to be elucidated.

HIV Gag-CA proteins play indispensable roles at various steps in the viral replication cycle (Freed and Martin, 2013). Our results described here clearly indicated that CA proteins of HIV-1 and HIV-2 are biochemically distinct. The data on the thermal stability of HIV-1/HIV-2 CA NTD proteins (**Figures 6**, **7**) may account for the unique polymerization properties of HIV-1 and HIV-2 CA proteins (**Figures 2**, **4**). As for the biological implications of our findings, we noticed a report showing that the HIV-2 GH123 virus carrying an identical Gag-CA with GL-AN exhibits faster uncoating (CA disassembly) kinetics in infected cells relative to the HIV-1 NL4-3 virus (Takeda et al., 2015). This observation is in good agreement with our results, supporting the notion that the instability/stability of CA proteins may affect the early viral replication phase (uncoating) of HIV-1 and HIV-2. The plausibility of our concept needs to be biochemically and biologically verified.

It is reasonable to consider Gag-CA as a therapeutic target, and indeed, there have been numerous attempts in this direction (Tang et al., 2003; Zhang et al., 2008; Tian et al., 2009; Blair et al., 2010; Jin et al., 2010; Curreli et al., 2011; Shi et al., 2011; Kortagere et al., 2012; Lemke et al., 2012; Lamorte et al., 2013; Matreyek et al., 2013; Bhattacharya et al., 2014; Fricke et al., 2014; Peng et al., 2014; Price et al., 2014). However, to the best of our knowledge, none of the anti-viral inhibitors described have proceeded to the steps to study their practical/clinical use. In the present study, we have demonstrated the close association of CA polymerization property *in vitro* and thermal stability of CA NTD as monitored by DSF. Thermal stability can be readily evaluated in large numbers simultaneously by a real-time PCR machine. It is thus practical to identify compounds by DSF that unusually destabilize or stabilize Gag-CA NTD proteins of HIV-1/HIV-2. In fact, a small molecule named PF74, previously reported to bind to HIV-1 CA and induce premature HIV-1 uncoating (Shi et al., 2011), was found to aberrantly increase the stability of NLCA NTD as revealed by DSF in our pilot experiments (manuscript in preparation). The system based on DSF (**Figures 6**, **7**) represents a promising new high-throughput screening method to search for durable and effective anti-HIV CA antivirals from a large library of candidate molecules.

AUTHOR CONTRIBUTIONS

YS, AA, and MN designed the research project. YS, AM, and TU performed the experiments. YS, AM, ND, TK, TU, AA, and MN discussed the results obtained. YS, TK, AA, and MN wrote the manuscript. All authors approved its submission.

FUNDING

This study was supported in part by a grant to MN from Japan Agency for Medical Research and Development, AMED (Research Program on HIV/AIDS: e-Rad ID number, 16768720).

ACKNOWLEDGMENT

We thank Ms. Kazuko Yoshida (Department of Microbiology, Tokushima University Graduate School of Medical Sciences, Tokushima, Japan) for editorial assistance.

REFERENCES

Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A., et al. (1986). Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* 59, 284–291.

Alfadhli, A., Dhenub, T. C., Still, A., and Barklis, E. (2005). Analysis of human immunodeficiency virus type 1 Gag dimerization-induced assembly. *J. Virol.* 79, 14498–14506. doi: 10.1128/JVI.79.23.14498-14506.2005

Barklis, E., Alfadhli, A., McQuaw, C., Yalamuri, S., Still, A., Barklis, R. L., et al. (2009). Characterization of the in vitro HIV-1 capsid assembly pathway. *J. Mol. Biol.* 387, 376–389. doi: 10.1016/j.jmb.2009.01.058

Bhattacharya, A., Alam, S. L., Fricke, T., Zadrozny, K., Sedzicki, J., Taylor, A. B., et al. (2014). Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci. U.S.A.* 111, 18625–18630. doi: 10.1073/pnas.1419945112

Black, L. R., and Aiken, C. (2010). TRIM5alpha disrupts the structure of assembled HIV-1 capsid complexes in vitro. *J. Virol.* 84, 6564–6569. doi: 10.1128/JVI.00210-10

Blair, W. S., Pickford, C., Irving, S. L., Brown, D. G., Anderson, M., Bazin, R., et al. (2010). HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog.* 6:e1001220. doi: 10.1371/journal.ppat1001220

Braaten, D., Aberham, C., Franke, E. K., Yin, L., Phares, W., and Luban, J. (1996a). Cyclosporine A-resistant human immunodeficiency virus type 1 mutants demonstrate that Gag encodes the functional target of cyclophilin A. *J. Virol.* 70, 5170–5176.

Braaten, D., Franke, E. K., and Luban, J. (1996b). Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *J. Virol.* 70, 3551–3560.

Byeon, I. J., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., et al. (2009). Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* 139, 780–790. doi: 10.1016/j.cell.2009.10.010

Campbell, S., and Vogt, V. M. (1995). Self-assembly in vitro of purified CA-NC proteins from *Rous sarcoma virus* and human immunodeficiency virus type 1. *J. Virol.* 69, 6487–6497.

Chen, B., and Tycko, R. (2010). Structural and dynamical characterization of tubular HIV-1 capsid protein assemblies by solid state nuclear magnetic resonance and electron microscopy. *Protein Sci.* 19, 716–730. doi: 10.1002/pro.348

Cortines, J. R., Lima, L. M. T. R., Mohana-Borges, R., Millen, T., de, A., Gaspar, L. P., et al. (2015). Structural insights into the stabilization of the human immunodeficiency virus type 1 capsid protein by the cyclophilin-binding domain and implications on the virus cycle. *Biochim. Biophys. Acta* 1854, 341–348. doi: 10.1016/j.bbapap.2014.12.008

Cortines, J. R., Monroe, E. B., Kang, S., and Prevelige, P. E. Jr. (2011). A retroviral chimeric capsid protein reveals the role of the N-terminal beta-hairpin in mature core assembly. *J. Mol. Biol.* 410, 641–652. doi: 10.1016/j.jmb.2011.03.052

Curreli, F., Zhang, H., Zhang, X., Pyatkin, I., Victor, Z., Altieri, A., et al. (2011). Virtual screening based identification of novel small-molecule inhibitors targeted to the HIV-1 capsid. *Bioorg. Med. Chem.* 19, 77–90. doi: 10.1016/j.bmc.2010.11.045

del Alamo, M., Rivas, G., and Mateu, M. G. (2005). Effect of macromolecular crowding agents on human immunodeficiency virus type 1 capsid protein assembly in vitro. *J. Virol.* 79, 14271–14281. doi: 10.1128/JVI.79.22.14271-14281.2005

Du, S., Betts, L., Yang, R., Shi, H., Concel, J., Ahn, J., et al. (2011). Structure of the HIV-1 full-length capsid protein in a conformationally trapped unassembled state induced by small-molecule binding. *J. Mol. Biol.* 406, 371–386. doi: 10.1016/j.jmb.2010.11.027

Ehrlich, L. S., Agresta, B. E., and Carter, C. A. (1992). Assembly of recombinant human immunodeficiency virus type 1 capsid protein in vitro. *J. Virol.* 66, 4874–4883.

Ehrlich, L. S., Liu, T., Scarlata, S., Chu, B., and Carter, C. A. (2001). HIV-1 capsid protein forms spherical (immature-like) and tubular (mature-like) particles in vitro: structure switching by pH-induced conformational changes. *Biophys. J.* 81, 586–594. doi: 10.1016/S0006-3495(01)75725-6

Fedorov, O., Niesen, F. H., and Knapp, S. (2012). Kinase inhibitor selectivity profiling using differential scanning fluorimetry. *Methods Mol. Biol.* 795, 109–118. doi: 10.1007/978-1-61779-337-0_7

Forshey, B. M., Shi, J., and Aiken, C. (2005). Structural requirements for recognition of the human immunodeficiency virus type 1 core during host restriction in owl monkey cells. *J. Virol.* 79, 869–875. doi: 10.1128/JVI.79.2.869-875.2005

Forshey, B. M., von Schwedler, U., Sundquist, W. I., and Aiken, C. (2002). Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol.* 76, 5667–5677. doi: 10.1128/JVI.76.11.5667-5677.2002

Freed, E. O., and Martin, M. A. (2013). "Human immunodeficiency viruses: replication," in *Fields Virology*, eds D. M. Knipe, P. M. Howley, J. I. Cohen, D. E. Griffin, R. A. Lamb, M. A. Martin, et al. (Philadelphia, PA: Lippincott Williams & Wilkins), 1502–1560.

Fricke, T., Buffone, C., Opp, S., Valle-Casuso, J., and Diaz-Griffero, F. (2014). BI-2 destabilizes HIV-1 cores during infection and prevents binding of CPSF6 to the HIV-1 capsid. *Retrovirology* 11:120. doi: 10.1186/s12977-014-0120-x

Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., et al. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* 87, 1285–1294. doi: 10.1016/S0092-8674(00)81823-1

Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., et al. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* 278, 849–853. doi: 10.1126/science.278.5339.849

Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science* 283, 80–83. doi: 10.1126/science.283.5398.80

Ganser-Pornillos, B. K., Cheng, A., and Yeager, M. (2007). Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell* 131, 70–79. doi: 10.1016/j.cell.2007.08.018

Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C., and Sundquist, W. I. (2004). Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol.* 78, 2545–2552. doi: 10.1128/JVI.78.5.2545-2552.2004

Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* 273, 231–235. doi: 10.1126/science.273.5272.231

Goff, S. P. (2013). "Retroviridae," in *Fields Virology*, eds D. M. Knipe, P. M. Howley, J. I. Cohen, D. E. Griffin, R. A. Lamb, M. A. Martin, et al. (Philadelphia, PA: Lippincott Williams & Wilkins), 1424–1473.

Goujon, C., Moncorge, O., Bauby, H., Doyle, T., Ward, C. C., Schaller, T., et al. (2013). Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* 502, 559–562. doi: 10.1038/nature12542

Grattinger, M., Hohenberg, H., Thomas, D., Wilk, T., Muller, B., and Kräusslich, H. G. (1999). In vitro assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A. *Virology* 257, 247–260. doi: 10.1006/viro.1999.9668

Gres, A. T., Kirby, K. A., KewalRamani, V. N., Tanner, J. J., Pornillos, O., and Sarafianos, S. G. (2015). Structural virology. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* 349, 99–103. doi: 10.1126/science.aaa5936

Grime, J. M., and Voth, G. A. (2012). Early stages of the HIV-1 capsid protein lattice formation. *Biophys. J.* 103, 1774–1783. doi: 10.1016/j.bpj.2012.09.007

Gross, I., Hohenberg, H., and Krausslich, H. G. (1997). In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur. J. Biochem.* 249, 592–600. doi: 10.1111/j.1432-1033.1997.t01-1-00592.x

Gross, I., Hohenberg, H., Wilk, T., Wiegers, K., Grattinger, M., Müller, B., et al. (2000). A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* 19, 103–113. doi: 10.1093/emboj/19.1.103

Guth, C. A., and Sodroski, J. (2014). Contribution of PDZD8 to stabilization of the human immunodeficiency virus type 1 capsid. *J. Virol.* 88, 4612–4623. doi: 10.1128/JVI.02945-13

Henning, M. S., Morham, S. G., Goff, S. P., and Naghavi, M. H. (2010). PDZD8 is a novel Gag-interacting factor that promotes retroviral infection. *J. Virol.* 84, 8990–8995. doi: 10.1128/JVI.00843-10

Jin, Y., Tan, Z., He, M., Tian, B., Tang, S., Hewlett, I., et al. (2010). SAR and molecular mechanism study of novel acylhydrazone compounds targeting HIV CA. *Bioorg. Med. Chem.* 18, 2135–2140. doi: 10.1016/j.bmc.2010.02.003

Kane, M., Yadav, S. S., Bitzegeio, J., Kutluay, S. B., Zang, T., Wilson, S. J., et al. (2013). MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* 502, 563–566. doi: 10.1038/nature12653

Kawamura, M., Sakai, H., and Adachi, A. (1994). Human immunodeficiency virus Vpx is required for the early phase of replication in peripheral blood mononuclear cells. *Microbiol. Immunol.* 38, 871–878. doi: 10.1111/j.1348-0421.1994.tb02140.x

Kortagere, S., Madani, N., Mankowski, M. K., Schon, A., Zentner, I., Swaminathan, G., et al. (2012). Inhibiting early-stage events in HIV-1 replication by small-molecule targeting of the HIV-1 capsid. *J. Virol.* 86, 8472–8481. doi: 10.1128/JVI.05006-11

Lamorte, L., Titolo, S., Lemke, C. T., Goudreau, N., Mercier, J. F., Wardrop, E., et al. (2013). Discovery of novel small-molecule HIV-1 replication inhibitors that stabilize capsid complexes. *Antimicrob. Agents Chemother.* 57, 4622–4631. doi: 10.1128/AAC.00985-13

Lampel, A., Bram, Y., Levy-Sakin, M., Bacharach, E., and Gazit, E. (2013). The effect of chemical chaperones on the assembly and stability of HIV-1 capsid protein. *PLoS ONE* 8:e60867. doi: 10.1371/journal.pone.0060867

Lanman, J., Sexton, J., Sakalian, M., and Prevelige, P. E. Jr. (2002). Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly in vitro. *J. Virol.* 76, 6900–6908. doi: 10.1128/JVI.76.14.6900-6908.2002

Lee, K., Ambrose, Z., Martin, T. D., Oztop, I., Mulky, A., Julias, J. G., et al. (2010). Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* 7, 221–233. doi: 10.1016/j.chom.2010.02.007

Lemke, C. T., Titolo, S., von Schwedler, U., Goudreau, N., Mercier, J. F., Wardrop, E., et al. (2012). Distinct effects of HIV-1 capsid assembly inhibitor families that bind the same site within the N-terminal domain of the viral CA protein. *J. Virol.* 86, 6643–6655. doi: 10.1128/JVI.00493-12

Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* 407, 409–413. doi: 10.1038/35030177

Lidon-Moya, M. C., Barrera, F. N., Bueno, M., Perez-Jimenez, R., Sancho, J., Mateu, M. G., et al. (2005). An extensive thermodynamic characterization of the dimerization domain of the HIV-1 capsid protein. *Protein Sci.* 14, 2387–2404. doi: 10.1110/ps.041324305

Liu, Z., Pan, Q., Ding, S., Qian, J., Xu, F., Zhou, J., et al. (2013). The interferon-inducible MxB protein inhibits HIV-1 infection. *Cell Host Microbe* 14, 398–410. doi: 10.1016/j.chom.2013.08.015

Mateu, M. G. (2002). Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus-1 capsid protein. *J. Mol. Biol.* 318, 519–531. doi: 10.1016/S0022-2836(02)00091-8

Matreyek, K. A., Yucel, S. S., Li, X., and Engelman, A. (2013). Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. *PLoS Pathog.* 9:e1003693. doi: 10.1371/journal.ppat.1003693

Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., et al. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nat. Struct. Biol.* 3, 763–770. doi: 10.1038/nsb0996-763

Morikawa, Y., Goto, T., and Momose, F. (2004). Human immunodeficiency virus type 1 Gag assembly through assembly intermediates. *J. Biol. Chem.* 279, 31964–31972. doi: 10.1074/jbc.M313432200

Niesen, F. H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* 2, 2212–2221. doi: 10.1038/nprot.2007.321

Peng, K., Muranyi, W., Glass, B., Laketa, V., Yant, S. R., Tsai, L., et al. (2014). Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *Elife* 3:e04114. doi: 10.7554/eLife.04114

Piroozmand, A., Khamsri, B., Fujita, M., Adachi, A., and Uchiyama, T. (2006). Morphological study on biologically distinct vpx/vpr mutants of HIV-2. *J. Med. Invest.* 53, 271–276. doi: 10.2152/jmi.53.271

Pornillos, O., Ganser-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., et al. (2009). X-ray structures of the hexameric building block of the HIV capsid. *Cell* 137, 1282–1292. doi: 10.1016/j.cell.2009.04.063

Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011). Atomic-level modelling of the HIV capsid. *Nature* 469, 424–427. doi: 10.1038/nature09640

Price, A. J., Fletcher, A. J., Schaller, T., Elliott, T., Lee, K., KewalRamani, V. N., et al. (2012). CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. *PLoS Pathog.* 8:e1002896. doi: 10.1371/journal.ppat.1002896

Price, A. J., Jacques, D. A., McEwan, W. A., Fletcher, A. J., Essig, S., Chin, J. W., et al. (2014). Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog.* 10:e1004459. doi: 10.1371/journal.ppat.1004459

Price, A. J., Marzetta, F., Lammers, M., Ylinen, L. M., Schaller, T., Wilson, S. J., et al. (2009). Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nat. Struct. Mol. Biol.* 16, 1036–1042. doi: 10.1038/nsmb.1667

Robinson, B. A., Reed, J. C., Geary, C. D., Swain, J. V., and Lingappa, J. R. (2014). A temporospatial map that defines specific steps at which critical surfaces in the Gag MA and CA domains act during immature HIV-1 capsid assembly in cells. *J. Virol.* 88, 5718–5741. doi: 10.1128/JVI.03609-13

Sakaguchi, T., Uchiyama, T., Huang, C., Fukuhara, N., Kiyotani, K., Nagai, Y., et al. (2002). Alteration of *Sendai virus* morphogenesis and nucleocapsid incorporation due to mutation of cysteine residues of the matrix protein. *J. Virol.* 76, 1682–1690. doi: 10.1128/JVI.76.4.1682-1690.2002

Sebastian, S., and Luban, J. (2005). TRIM5alpha selectively binds a restriction-sensitive retroviral capsid. *Retrovirology* 2:40. doi: 10.1186/1742-4690-2-40

Shi, J., Zhou, J., Shah, V. B., Aiken, C., and Whitby, K. (2011). Small-molecule inhibition of human immunodeficiency virus type 1 infection by virus capsid destabilization. *J. Virol.* 85, 542–549. doi: 10.1128/JVI.01406-10

Shibata, R., Miura, T., Hayami, M., Ogawa, K., Sakai, H., Kiyomasu, T., et al. (1990). Mutational analysis of the human immunodeficiency virus type 2 (HIV-2) genome in relation to HIV-1 and simian immunodeficiency virus SIV (AGM). *J. Virol.* 64, 742–747.

Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., et al. (2006). Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5514–5519. doi: 10.1073/pnas.0509996103

Takeda, E., Kono, K., Hulme, A. E., Hope, T. J., Nakayama, E. E., and Shioda, T. (2015). Fluorescent image analysis of HIV-1 and HIV-2 uncoating kinetics in the presence of Old World monkey TRIM5α. *PLoS ONE* 10:e0121199 doi: 10.1371/journal.pone.0121199

Tang, C., Loeliger, E., Kinde, I., Kyere, S., Mayo, K., Barklis, E., et al. (2003). Antiviral inhibition of the HIV-1 capsid protein. *J. Mol. Biol.* 327, 1013–1020. doi: 10.1016/S0022-2836(03)00289-4

Ternois, F., Sticht, J., Duquerroy, S., Krausslich, H. G., and Rey, F. A. (2005). The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat. Struct. Mol. Biol.* 12, 678–682. doi: 10.1038/nsmb967

Tian, B., He, M., Tang, S., Hewlett, I., Tan, Z., Li, J., et al. (2009). Synthesis and antiviral activities of novel acylhydrazone derivatives targeting HIV-1 capsid protein. *Bioorg. Med. Chem. Lett.* 19, 2162–2167. doi: 10.1016/j.bmcl.2009.02.116

Tsiang, M., Niedziela-Majka, A., Hung, M., Jin, D., Hu, E., Yant, S., et al. (2012). A trimer of dimers is the basic building block for human immunodeficiency virus-1 capsid assembly. *Biochemistry* 51, 4416–4428. doi: 10.1021/bi300052h

von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R., et al. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* 17, 1555–1568. doi: 10.1093/emboj/17.6.1555

von Schwedler, U. K., Stray, K. M., Garrus, J. E., and Sundquist, W. I. (2003). Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J. Virol.* 77, 5439–5450. doi: 10.1128/JVI.77.9.5439-5450.2003

Yoo, S., Myszka, D. G., Yeh, C., McMurray, M., Hill, C. P., and Sundquist, W. I. (1997). Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *J. Mol. Biol.* 269, 780–795. doi: 10.1006/jmbi.1997.1051

Zhang, H., Zhao, Q., Bhattacharya, S., Waheed, A. A., Tong, X., Hong, A., et al. (2008). A cell-penetrating helical peptide as a potential HIV-1 inhibitor. *J. Mol. Biol.* 378, 565–580. doi: 10.1016/j.jmb.2008.02.066

Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., et al. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* 497, 643–646. doi: 10.1038/nature12162

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Miyazaki, Miyake, Doi, Koma, Uchiyama, Adachi and Nomaguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 May 2016
doi: 10.3389/fmicb.2016.00635

Pathogenicity, Transmission and Antigenic Variation of H5N1 Highly Pathogenic Avian Influenza Viruses

Peirong Jiao[1,2,3,4†], *Hui Song*[1,2,3,4†], *Xiaoke Liu*[1,5], *Yafen Song*[1,2,3,4], *Jin Cui*[2,3,4], *Siyu Wu*[1,2,3,4], *Jiaqi Ye*[1,2,3,4], *Nanan Qu*[1,2,3,4], *Tiemin Zhang*[6] *and Ming Liao*[1,2,3,4*]

[1] National and Regional Joint Engineering Laboratory for Medicament of Zoonosis Prevention and Control, Guangzhou, China, [2] Key Laboratory of Animal Vaccine Development, Ministry of Agriculture, Guangzhou, China, [3] Key Laboratory of Zoonosis Prevention and Control of Guangdong, Guangzhou, China, [4] College of Veterinary Medicine, South China Agricultural University, Guangzhou, China, [5] Pulike Biological Engineering Inc., Luoyang, China, [6] College of Engineering, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University Graduate School, Japan

Reviewed by:
Huaguang Lu,
The Pennsylvania State University, USA
Eri Nobusawa,
National Institute of Infectious Diseases, Japan

***Correspondence:**
Ming Liao
mliao@scau.edu.cn

[†]*These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *09 November 2015*
Accepted: *18 April 2016*
Published: *06 May 2016*

Citation:
Jiao P, Song H, Liu X, Song Y, Cui J, Wu S, Ye J, Qu N, Zhang T and Liao M (2016) Pathogenicity, Transmission and Antigenic Variation of H5N1 Highly Pathogenic Avian Influenza Viruses. Front. Microbiol. 7:635. doi: 10.3389/fmicb.2016.00635

H5N1 highly pathogenic avian influenza (HPAI) was one of the most important avian diseases in poultry production of China, especially in Guangdong province. In recent years, new H5N1 highly pathogenic avian influenza viruses (HPAIV) still emerged constantly, although all poultry in China were immunized with H5N1 vaccinations compulsorily. To better understand the pathogenicity and transmission of dominant clades of the H5N1 HPAIVs in chicken from Guangdong in 2012, we chose a clade 7.2 avian influenza virus named A/Chicken/China/G2/2012(H5N1) (G2) and a clade 2.3.2.1 avian influenza virus named A/Duck/China/G3/2012(H5N1) (G3) in our study. Our results showed that the chickens inoculated with 10^3 EID_{50} of G2 or G3 viruses all died, and the titers of virus replication detected in several visceral organs were high but different. In the naive contact groups, virus shedding was not detected in G2 group and all chickens survived, but virus shedding was detected in G3 group and all chickens died. These results showed that the two clades of H5N1 HPAIVs had high pathogenicity in chickens and the contact transmission of them was different in chickens. The results of cross reactive HI assay showed that antigens of G2 and G3 were very different from those of current commercial vaccines isolates (Re-4, Re-6, and D7). And to evaluate the protective efficacy of three vaccines against most isolates form Guangdong belonging to clade 2.3.2.1 in 2012, G3 was chosen to challenge the three vaccines such as Re-4, Re-6, and D7. First, chickens were immunized with 0.3 ml Re-4, Re-6, and D7 inactivated vaccines by intramuscular injection, respectively, and then challenged with 10^6 EID_{50} of G3 on day 28 post-vaccination. The D7 vaccine had 100% protection against G3 for chickens, the Re-6 vaccine had 88.9%, and the Re-4 vaccine only had 66.7%. Our results suggested that the D7 vaccine could prevent and control H5N1 virus outbreaks more effectively in Guangdong. From the above, it was necessary to conduct continuously epidemiological survey and study the pathogenicity and antigenic variation of avian influenza in Southern China.

Keywords: H5N1, pathogenicity, transmissibility, antigenic variation, vaccine

INTRODUCTION

Influenza A viruses are the most important pathogen of three types (A, B, and C) of influenza viruses, to both the poultry industry and human health. To date, avian influenza viruses representing 16 HA and 9 NA subtypes have been detected in wild birds and poultry throughout the world (Webster et al., 1992; Fouchier and Munster, 2009). According to the virulence of viruses, avian influenza viruses are divided into highly pathogenic avian influenza virus (HPAIV), low pathogenic avian influenza virus (LPAIV), and non-pathogenic avian influenza virus (NPAIV). However, only some influenza viruses of H5 and H7 subtype are highly pathogenic to poultry.

The H5N1 HPAIV was first isolated in Guangdong, China in 1996 (Xu et al., 1999). In 1997, H5N1 HPAIVs had repeatedly caused serious outbreaks among poultry farms and markets in Hong Kong, which resulted in heavy losses. And it was the first report that H5N1 HPAIV infected human in "Hong Kong Flu" in 1997, causing six deaths in 18 infection cases (Claas et al., 1998; Subbarao et al., 1998). In 2002, a new H5N1 HPAI outbreak in Hong Kong infected millions of birds, including several types of wild water fowl. This was the first time the H5N1 HPAIV was found to infect water fowl (Lee et al., 2005; Nguyen et al., 2005). Between 2003 and 2005, the H5N1 HPAI repetitively broke out in East Asia and South Asia, and even spread to Europe and Africa. This resulted in more than 150 million birds dead or slaughtered and 53 human fatalities (Sturm-Ramirez et al., 2004; World Health Organization, 2005). Since 2003, the H5N1 HPAI had continued influencing more than 60 states or areas including Laos, Vietnam, Thailand, Hong Kong, and China (World Health Organization, 2013b). From 2003 to 1 May 2015, 840 laboratory-confirmed human cases of H5N1 HPAIVs infection were officially reported to WHO from 16 countries; of these cases, 447 died (World Health Organization, 2015). Consequently, H5N1 HPAIVs are zoonotic etiological agents recognized as a severe threat to both the poultry industry and human public health around the world.

In terms of antigenic characteristics, H5N1 HPAIVs were divided into 10 clades (0–9) and numerous subclades by World Health Organization/World Organization for Animal Health/Food and Agriculture Organization H5N1 Evolution Working Group (2008)[1]. Complicated breeding environments, the long distance transport of live poultry, and wild bird migration resulted in all known clades circulating endlessly in poultry in China. Especially during 2005 and 2006, H5N1 viruses of clades 2.2, 2.3.2, 2.3.4, 4, 7, and 9 circulated all over China. Since 2007, viruses of clades 2.3.2, 2.3.4, and 7 have predominantly co-circulated continuously in domestic poultry and waterfowl in China (Smith et al., 2009; Jiang et al., 2010; Li et al., 2010). Later, several studies results showed that the pathogenicity of clade 2.3.2 viruses were intensifying in aquatic birds (Sakoda et al., 2010). Viruses of clade 7 began spreading in chickens across the northern of China in 2005, which had a high pathogenicity in chickens, but only a few viruses were isolated from aquatic birds. In 2008, H5N1 HPAI caused clade 7.2 viruses broke out in several cities in North China and caused a considerable amount of deaths in poultry. In 2010, a new H5N1 HPAIV belonging to clade 2.3.2.1 was isolated from South Asia, and 48 humans were reported to have been infected with the virus (Reid et al., 2011). In 2011, the H5N1 HPAI caused clade 2.3.2.1 viruses broke out in crows (World Health Organization, 2012). From 2012 to 2013, the H5N1 HPAIVs belonging to 2.3.2.1, 2.3.4, and 7.2 clades were detected in birds and/or environmental samples in China (World Health Organization, 2013a), but the most isolates belonged to clade 2.3.2.1. The pathogenicity of different clades varied in poultry and wild birds, but the movement and interaction of H5N1 viruses between them was still not clear until now.

To better understand the pathogenicity and transmissibility of different clade of H5N1 isolates from poultry in Guangdong in 2012, we selected two viruses—A/Chicken/China/G2/2012(H5N1) (G2) and A/Duck/China/G3/2012(H5N1) (G3)—to carry out their infection experiments. To evaluate the antigenic variation of these viruses and protective efficacy of current commercial vaccines against most isolates from Guangdong in 2012, G3 (belonging to clade 2.3.2.1) was chosen to challenge three commercial vaccines such as Re-4, Re-6, and D7.

MATERIALS AND METHODS

H5N1 HPAIV Variants and Propagation

The two H5N1 HPAIVs—A/Chicken/China/G2/2012(H5N1) (G2) and A/Duck/China/G3/2012(H5N1) (G3)—used in this study were isolated from cloacal swabs of apparently healthy birds in live bird markets during 2012. They were purified and propagated by three rounds of limiting dilution in the allantoic cavity of 9–11 days old specific-pathogen-free (SPF) embryonated chicken eggs (Jiao et al., 2014; Yuan et al., 2014). The allantoic fluid from multiple eggs was pooled, clarified by centrifugation, and frozen in aliquots at −70°C. The G2 and G3 inactivated antigens and positive serums were provided by College of Veterinary Medicine, South China Agricultural University. The 50% egg infectious dose (EID_{50}) was calculated according to the method published by Reed and Muench (1938) using the serial titration of eggs. All experiments were carried out in Animal Biosafety Level 3 (ABSL-3) facilities.

Genetic and Phylogenetic Analyses

The viral RNA was extracted from the allantoic fluid supernatant using Trizol LS Reagent (Life Technologies, Inc.). A reverse transcription polymerase chain reaction (RT-PCR) was conducted using Superscript III (Invitrogen, Carlsbad, CA, USA) and Uni12 (5-AGCAAAAGCAGG-3) primer. Eight genes were amplified using universal primers (Hoffmann et al., 2001), and the PCR products were purified using the mini PCR Purification Kit (Promega). Sequencing was performed by Shanghai Invitrogen Biotechnology Co., Ltd. The sequencing

[1]Food and Agricultural Organization(FAO). H5N1 HPAI Global Overview - July and August 2010, prepared by EMPRESS/GLEW, Issues No. 24.

data were compiled with the Seqman program of Lasergene 7 (DNASTAR, Inc.). Amino acid sequence similarities were identified with the Lasergene 7 Megalign program (DNASTAR). The hemagglutinin (HA) gene phylogenetic tree of the H5N1 HPAIVs was created with MEGA 5 software (Sinauer Associates, Inc., Sunderland, MA).

The nucleotide sequences of A/Chicken/China/G2/2012(H5N1) (G2) and A/Duck/China/G3/2012(H5N1) (G3) were available from GenBank under the accession numbers KU851866-KU851867.

Pathogenicity and Transmission

Five-week-old SPF White Leghorn chickens were purchased from Beijing Merial Vital Laboratory Animal Technologies Co., LTD, Beijing, China.

To determine the pathogenicity and transmission of the two H5N1 HPAIVs, twenty-seven chickens were equally divided into three groups G2, G3, and control. Six chickens of G2 and G3 group were inoculated intranasally with 10^3 EID_{50} of G2 or G3 viruses, respectively; the other three chickens of each group were inoculated intranasally with the same volume of phosphate buffered saline (PBS), as naive contact housed with the inoculated chickens. The chickens of control group were inoculated intranasally with the same volume of PBS. All chickens were observed for clinical symptoms for 14 days. Three inoculated chickens in each group were euthanized at 3 days post-inoculation (DPI), and the lungs, kidneys, liver, heart, spleen, and brain were collected. Similar executions were performed on chickens that died during the observation. Oropharyngeal and cloacal swabs were collected from all chickens at 3, 5, 7, 9, and 11 DPI, and suspended in 1 ml isolation media PBS (pH 7.4). All of the tissues and swabs were collected and titrated for virus infectivity in eggs, as described previously (Chen et al., 2004; Jiao et al., 2008). Seroconversion of the surviving chickens on 14 DPI was confirmed by hemagglutinin inhibition (HI) test. HI titers of the serums were detected using 1% chicken red blood cells by a standard method (Takatsy and Barb, 1973). All animal experiments were conducted under the guidance of SCAU's Institutional Animal Care and Use Committee. Our animal experiments in this study had been approved by SCAU and were carried out in high-efficiency particulate air-filtered isolators (size: 2200 × 850 × 1700 mm) and ABSL-3 facilities.

Vaccine-Challenge

To evaluate the antigenic variation of these viruses and protective efficacy of current commercial vaccines against most isolates from Guangdong in 2012, 3-week-old SPF White Leghorn chickens were purchased from Beijing Merial Vital Laboratory Animal Technologies Co., LTD, Beijing, China. Re-4 and Re-6 vaccines strain inactivated antigens, positive serums, and vaccines were purchased from Weike Biotechnology Co., Ltd., Harbin, China. D7 (H5N2) vaccines strain inactivated antigens, positive serums, and vaccines were purchased from Guangzhou South China Biological Medicine Co. Ltd., Guangdong, China.

Thirty-six chickens were divided into four groups ($n = 9$), and three groups were immunized with 0.3 ml of Re-4, Re-6, or D7 inactivated vaccines via intramuscular injection, the control group received 0.3 ml of PBS intramuscularly. Serum was collected from every chicken on 14 and 28 day-post-vaccination (DPV) for HI titers determination.

At 28 DPV, chickens were intranasally challenged with 200ul 10^6EID_{50} of A/Duck/China/G3/2012(H5N1) (G3). Oropharyngeal and cloacal swabs were taken on days 3, 5, 7, 9, and 11 post-challenge, including chickens that died during this period. All swabs were immediately suspended in 1 ml isolation media PBS, which were inoculated into 9–10 days old embryonated chicken eggs for examination of virus shedding. All surviving chickens were observed for clinical symptoms for 14 days and collected serum for seroconversion detection in the end.

RESULTS

Genetic and Phylogenetic Analysis

The HA genes of each virus were sequenced to determine the molecular evolution of the two viruses. The sequences were compared with representative H5N1 sequences obtained from GenBank. According to antigenic characteristics by the WHO, the HA gene of G2 belonged to clade 7.2, and that of G3 belonged to clade 2.3.2.1 (**Figure 1**). Their HA genes had a series of basic amino acids at the cleavage site of the HA (-RRRKR/GLF-), which represents the high pathogenicity of the H5N1 AIVs in poultry (Gohrbandt et al., 2011).

The amino acid sequences of the two viruses revealed five conservative potential N-linked glycosylation sites in HA (26, 27, 39, 499, and 558): three in HA1 (26, 27, and 39) and two in

TABLE 1 | Cleavage site and potential glycosylation sites in HA of the two H5N1 HPAIVs (A/Chicken/China/G2/2012(H5N1) = G2 and A/Duck/China/G3/2012(H5N1) = G3).

Strains	Cleavage site	potential glycosylation sites										
	342-347	26	27	39	155	169	178	209	251	301	499	558
	-RRRKR/G-	NNS	NST	NVT	NSS	NNT	NNT	NLT	NDT	NSS	NGT	NGS
G2	+[a]	+	+	+	NPS	+	–[b]	–	+	+	+	+
G3	+	+	+	+	+	–	+	–	–	NYS	+	+

[a] *The "+" means the amino acid sequences of glycosylation sites are same with list above.*
[b] *The "–" means the glycosylation sites are lost.*

HA2 (499 and 558). In addition, the G2 virus lost two potential N-linked glycosylation sites in 178 (NNT) and 209 (NPT), and amino acids at 155 the glycosylation site changed from NSS to NPS (**Table 4**). The G3 HA lost three potential N-linked glycosylation sites in 169 (NNT), 209 (NPT) and 251 (NDT), and amino acids at 301 the glycosylation site changed from NSS to NYS (**Table 1**).

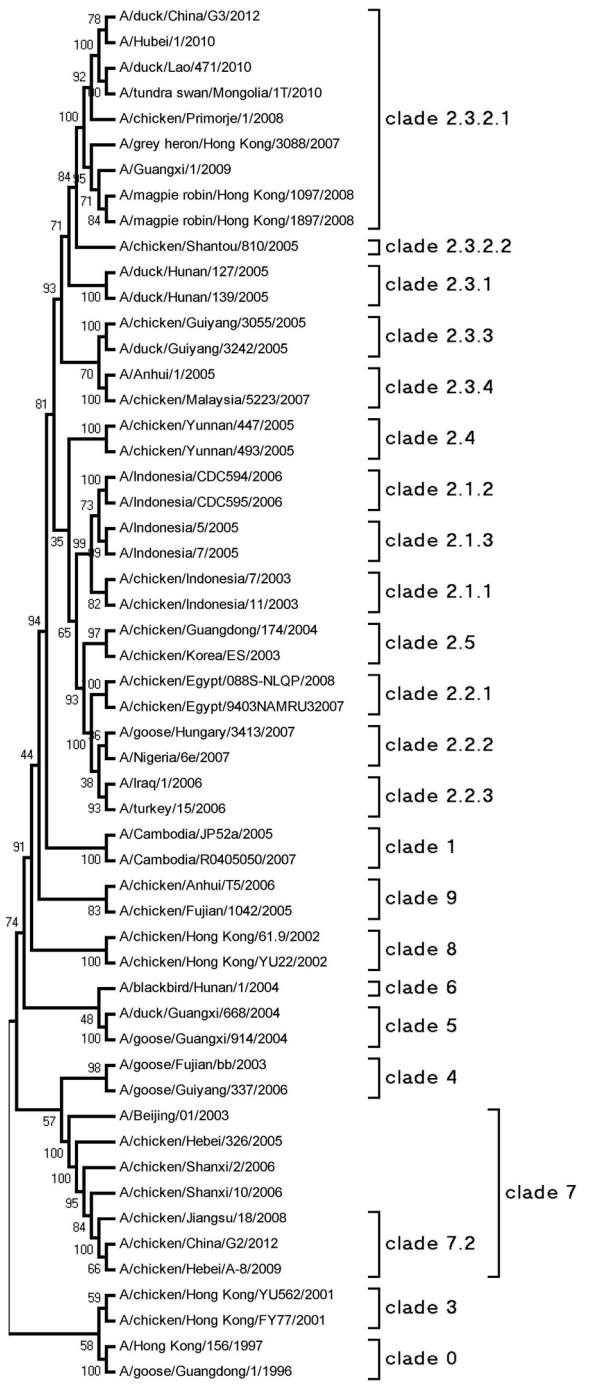


FIGURE 1 | Phylogenetic analysis of HA. The trees were constructed by using the neighbor joining method with the Maximum Composite Likelihood model and MEGA5 software with 1000 bootstrap replicates based on the following sequences: HA (A), nucleotides (nt) 29–1732.

Pathogenicity of H5N1 HPAIVs in Chickens

To evaluate the pathogenicity of the two H5N1 HPAIVs, six chickens of each group were inoculated intranasally with 100 μl 10^3 EID_{50} G2, G3, or PBS, respectively. All chickens in the G3 group began to show clinical typical symptoms as early as two DPI, and were dead by four DPI (**Figure 2**). However, the inoculated chickens of the G2 group showed clinical signs by four DPI, and all died by eight DPI (**Figure 2**). So the lethality of G2 and G3 viruses in chickens was 100% (**Table 2**).

Eyelid edema, insensibility, diminished appetite and thirst, roughened hair coats, comb cyanosis, torticollis, ataxia, and other neurological symptoms were observed among dead chickens infected with G2 and G3. At necropsy of chickens dead from infection we found slight petechial hemorrhaging in subcutaneous fat, hyperaemia, haemorrhagia, and nig-necrosis in the lungs; hepatomegaly and an amber liver; and hyperaemia and haemorrhagia in the stomachus glandularis. In short, both G2 and G3 viruses produced apparent clinical symptoms and typical pathological changes in severe infected chickens.

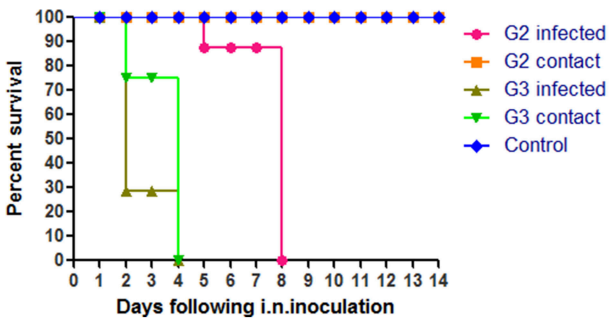


FIGURE 2 | Lethality of the G2 and G3 viruses in SPF chickens. The G2 infected chickens were inoculated intranasally with 100μl 10^3EID_{50} G2 viruses and the G2 contact chickens were housed with them without inoculate. The G3 infected chickens were inoculated intranasally with 100 ul 10^3EID_{50} G3 viruses and the G3 contact chickens were housed with them without inoculate.

TABLE 2 | Clinical situations and lethality of chickens after inoculated intranasally with the two H5N1 HPAIVs (A/Chicken/China/G2/2012 (H5N1) = G2 and A/Duck/China/G3/2012(H5N1) = G3).

Strains		Clinical symptom rates	Virus shedding proportions	Survival
G2	Infected[a]	100%	100%	0
	Contact[b]	100%	0	100%
G3	Infected	100%	100%	0
	Contact	100%	100%	0
Control		0	0	100%

[a] *Chickens inoculated with virus.*
[b] *Contact chickens housed with those inoculated.*

To evaluate the replication of the two viruses in chickens, three inoculated chickens in each group were euthanized at three DPI, and the lungs, kidneys, liver, heart, spleen, and brain were collected. Oropharyngeal and cloacal swabs were collected from chickens of each group at 3, 5, 7, 9, and 11 DPI. All of the tissues and swabs were collected and titrated for virus infectivity. In G2 inoculated chickens, the virus replicated in all tested organs on three DPI and the mean titers were 4.33 log EID_{50} in the heart, 2.92 log EID_{50} in the liver, 2.67 log EID_{50} in the spleen, 1.58 log EID_{50} in the lungs, 3.67 log EID_{50} in the kidneys, and 3.25 log EID_{50} in the brain (**Table 3**). In G3 virus inoculated chickens, the virus replicated to higher and the mean titers were 5.25 log EID_{50} in the liver and 5.5 log EID_{50} in the heart, spleen, lungs, kidneys, and brain, respectively (**Table 3**). Above all, the replication of G3 in chicken was much higher than that of G2.

In the G2 group, virus shedding was detected from the oropharynx and cloaca swabs in inoculated chickens within seven DPI. The virus titers from oropharynx swabs and cloacal swabs were 2.92 and 1.96 log EID_{50}, 2.63 and 1.88 log EID_{50}, and 2.63 and 1.63 log EID_{50} on three DPI, five DPI, and seven DPI, respectively (**Table 4**). All of the chickens in the G2 group died within eight DPI. G3 virus shedding could be tested from both oropharyngeal and cloacal swabs in inoculated chickens on three DPI, and the virus titers were all 4.5 logEID_{50} (**Table 4**). All of the chickens in the G3 group died within four DPI. These showed that the duration of virus shedding of chickens infected with G2 was 8 days and was longer than the 4 days of G3, but the titers of replication of G2 group were lower than that of G3. Therefore, our results indicated that both G2 and G3 were highly pathogenic to chickens, and the replication of the G3 virus was higher than that of the G2.

Transmission of H5N1 HPAIVs in Chickens

To understand the naive contact transmission of these two viruses, three SPF chickens were inoculated intranasally with 0.1 ml PBS as naive control group and housed with inoculated chickens of the G2 and G3 groups, respectively. Oropharyngeal and cloacal swabs were collected from them at 3, 5, 7, 9, and 11 DPI. All surviving chickens were observed for 14 days. We collected and titrated the tissues and swabs for virus infectivity.

During the observation period, the naive contact chickens in the G2 group began to show mild clinical signs, such as spirits atrophy and inappetence, by five DPI and these mild clinical signs disappeared by seven DPI, all chickens survived for 14 days and

TABLE 3 | Replication of the two H5N1 HPAIVs (A/Chicken/China/G2/2012(H5N1) = G2 and A/Duck/China/G3/2012(H5N1) = G3) in SPF chickens[a].

Strains		Virus titers in SPF chickens at 3 DPI ($log_{10}EID_{50}$/0.1ml) in[b]					
		Heart	Liver	Spleen	Lungs	Kidneys	Brain
G2	Infected	4.33 ± 0.52	2.92 ± 0.52	2.67 ± 0.14	1.58 ± 0.14	3.67 ± 0.52	3.25 ± 0.5
	Contact	0	0	0	0	0	0
G3	Infected	5.5 ± 0	5.25 ± 0.43	5.5 ± 0	5.5 ± 0	5.5 ± 0	5.5 ± 0
	Contact	5.5 ± 0	4.75 ± 0.66	5 ± 0.66	4.83 ± 0.63	5.25 ± 0.43	5.0 ± 0.43
Control		0	0	0	0	0	0

[a]Six SPF chickens were inoculated intranasally (i.n.) with 10^3 EID_{50} of virus in a 0.1 ml volume in G2 and G3 group, and three naive contact chickens housed with them, respectively; on 3 DPI, three inoculated and all dead naive chickens in each group were euthanized, and virus titers were determined in samples of heart, liver, spleen, lungs, kidneys, and brain in eggs.
[b]For statistical analysis, a value of 1.5 was assigned if the virus was not detected from the undiluted sample in three embryonated hen eggs (Sun et al., 2011). Virus titers are expressed as means ± standard deviation in log10EID_{50}/0.1 ml of tissue.

TABLE 4 | Virus titers in oropharyngeal and cloacal swabs from chickens after inoculated with the two H5N1 HPAIVs (A/Chicken/China/G2/2012(H5N1) = G2 and A/Duck/China/G3/2012(H5N1) = G3).

Strains		Days post-inoculation ($log_{10}EID^{50}$/0.1 ml) ± SD[a]							
		3 day		5 day		7 day		9 day	
		Oropharyngeal swabs	Cloacal swabs	Oropharyngeal swabs	Cloacal swabs	Oropharyngeal swabs	Cloacal swabs	Oropharyngeal swabs	Cloacal swabs
G2	Infected[b]	2.92 ± 1.11 (5/6)	1.96 ± 0.46 (5/6)	2.63 ± 0.18 (2/2)	1.88 ± 0.53 (1/2)	2.63 ± 0.18 (2/2)	1.63 ± 0.18 (1/2)	ND[d]	ND
	Contact[c]	0	0	0	0	0	0	0	0
G3	Infected	4.5 ± 0(1/1)	4.5 ± 0(1/1)	ND	–	–	–	–	–
	Contact	4.5 ± 0(2/2)	4.5 ± 0(2/2)	ND	–	–	–	–	–
Control		0	0	0	0	0	0	0	0

[a]For statistical purposes, a value of 1.5 was assigned if virus was not detected from the undiluted sample in three embryonated hen's eggs (Sun et al., 2011).
[b]Chickens inoculated with virus.
[c]Contact chickens housed with those inoculated.
[d]ND: not detected. Chickens all died.

the seroconversion rate was 100%. The naive contact chickens in the G3 group began to show clinical signs by two DPI and all died by four DPI (**Figure 2**). In the G3 naive contact group, virus replication titers were 5.5 log EID_{50} in the heart, 4.75 log EID_{50} in the liver, 5 log EID_{50} in the spleen, 4.83 log EID_{50} in the lungs, 5.25 log EID_{50} in the kidneys, and 5 log EID_{50} in the brain (**Table 3**). However, no virus was detected in tissue samples of the G2 naive contact group. All of the naive contact chickens in G3 group could shed virus from the oropharynx and cloaca swabs; the virus titers were both 4.5 log EID_{50} at three DPI, which was the same as that of inoculated chickens (**Table 4**). In the chickens of the G2 naive contact group, virus shedding could not be detected all along (**Table 4**). These showed that the lethality of naive contact chickens of G3 was 100% and that of G2 was 0 although both G2 and G3 had naive contact transmission in chickens (**Table 2**). So G3 virus had stronger transmissibility between chickens by naive contact than G2.

Antigenic Variation of HPAIV and Protective Efficacy of Current Commercial Vaccines

To characterize antigenic variation of the two H5N1 HPAIVs and the current commercially vaccines strains, we carried out the cross reactive HI assay. The cross reactive HI antibody titers of anti-Re-4, anti-Re-6, and anti-D7 serum in reaction with G2 antigen were 5 log2, 1 log2, and 2 log2, respectively; and those in reaction with G3 antigen were 5 log2, 6 log2, and 9 log2, respectively (**Table 5**). Our results showed that the antigens of G2 and G3 were very different from those of vaccine isolates. Therefore, we should assess the immunogenicity and effectiveness of three current commercially inactivated vaccines against these isolates in 2012 in Guangdong of China.

Because H5N1 HPAIVs of clade 2.3.2.1 have most isolates from Guangdong in 2012, we estimated the effectiveness of current commercial vaccines against G3. Three-week-old SPF chickens were immunized with inactivated vaccines such as Re-4, Re-6, and D7, respectively. Serum from every group was collected at 14 and 28 DPV for HI test, respectively. Then, all chickens at 28 DPV were challenged intranasally with 200 μl 10^6 EID_{50} of G3. In the Re-4 group, chickens were challenged with G3 when the mean HI titer was 9.4 log2 at 28 DPV. The chickens began to show clinical symptoms on three DPI and die on six DPI, shedding virus was tested on three to nine DPI, and the mortality and virus shedding proportion were 33.3 and 88.9%, respectively (**Table 6**). In the Re-6 group, chickens were challenged when the mean HI titer was 7.0 log2 at 28 DPV. The chickens began to die on six DPI, virus shedding was detected on three to eleven DPI, and the mortality and virus shedding proportion were 11.1 and 44.4%, respectively (**Table 6**). Those results showed that the Re-6 vaccine has a certain degree of protection against the G3 virus. In the D7 group, chickens were challenged when the mean HI titer was 7.4 log2 at 28 DPV, and no virus shedding or death was found during the observation period (**Table 6**). In the non-immunized control group, the HI titer was zero. All chickens were detected virus shedding and died during the observation period. These findings showed that the protection rate of the D7 vaccine against G3 was 100%, that of Re-6 was 88.9%, and Re-4 was 66.7%.

TABLE 5 | Cross reactive hemagglutination inhibition (HI)[a] antibody titers of anti-serum against five avian influenza virus antigens.

Strains	Clade	Serum of HI titers (log2)				
		Re-4	Re-6	D7	G2	G3
Re-4	7	10	3	8	2	5
Re-6	2.3.2	5	8	10	1	10
D7	2.3.2	5	6	9	1	9
G2	7.2	5	1	2	8	5
G3	2.3.2.1	5	6	9	0	10

[a] *The cross reactive HI assays were carried out according to WHO standard method.*

TABLE 6 | Results of hemagglutination inhibition (HI) titers from serum samples of chickens at 28 DPV. And the protection rates of three vaccines against G3 virus challanged[a].

Groups	Before challenged	Post-challenge	Survival rate (%)
	HI antibody titers[b] (log2)	Viruses shedding (%)	
Re-4	9.40[c]	88.9	66.7
Re-6	7.20	44.4	88.9
D7	7.50	0	100
Control	0	100	0

[a] *Thirty-six three-week-old SPF chickens were divided into four groups and immunized with inactivated Re-4, Re-6, D7, and PBS respectively. At 28 DPV, all chickens were challenged intranasally with 200 ul 10^6 EID_{50} of G3.*

[b] *Serum samples from Re-4 group, Re-6 group, and D7 group were detected with Re-4, Re-6, and D7 inactivated antigens, respectively. Serum samples from control group were detected simultaneously with Re-4, Re-6, and D7 inactivated antigens.*

[c] *Geometric mean titer (GMT).*

In a word, the mean HI titers in all immune groups were higher than 6 log2, which indicated that these three commercial vaccines (Re-4, Re-6, and D7) had good immunogenicity in chickens. And the results of challenge study showed that these vaccines gave certain protection against G3, but their protection rates were different. Combined with the results of the cross reactive HI assay, we found that some vaccine strains were not antigenically well-matched with epidemic isolates, so the protective effects of the three vaccines varied.

DISCUSSION

The first H5N1 HPAIV in China was isolated from sick geese in Guangdong province in 1996 (Xu et al., 1999). In the following years, H5N1 HPAIVs repeatedly caused serious outbreaks in South China, especially in Hong Kong, and resulted in heavy losses of economy and life. Most of H5N1 viruses rapidly spread and induced large numbers of death within 2 or 3 days in chickens. The ducks and geese infected H5N1 HPAIVs showed no clinical symptoms in the past, but new H5N1 HPAIVs could attacked ducks and/or geese and caused deaths in recent years (Li et al., 2010). In addition, more and more mammal was susceptible to H5N1 HPAIV by natural or laboratory infections. Felines, including cat, tiger, lion, leopard, clouded leopard, and Asiatic

golden cat were highly susceptible to H5N1 HPAIV (Reperant et al., 2009). The domestic dog, hamster, rhesus macaque, cynomolgus, palm civet, red fox and raccoon could be potentially fatal by H5N1 HPAIV. Pika, domesticated swine, cattle, donkey, rat, and rabbit can exhibit asymptomatic or nonfatal infections by H5N1 HPAIV (U.S. Geological Survey, 2011). From 2010 to 2013 the dominant clades of H5N1 HPAIVs co-circulating in South China were 2.3.2.1 and 7.2 although other clades, such as 2.3.4, had occasionally been detected (World Health Organization, 2013a). In our study, the G2 and G3 strains from poultry in Guangdong in 2012 belonged to clades 7.2 and 2.3.2.1, respectively. The inoculation dose was mostly 10^6 EID_{50} or 10^5 EID_{50} in previous pathogenicity studies, but here we selected a medium infective dose (10^3 EID_{50}) to observe the difference of pathogenicity and transmission of the H5N1 HPAIVs. This might be one of the reasons why the naive contact chickens in G2 group showed mild clinical symptoms without any viruses shedding or death. In our study, both G2 and G3 virus could highly replicated in the heart, liver, brain, spleen, kidneys, and lungs of infected chickens, virus shedding could be detected from all infected chickens during survival, and the lethal rates were both 100%. These results showed that G2 and G3 virus had high pathogenicity to chickens. By this token, the new H5N1 HPAIVs of these two clades in South China still had high pathogenicity to chickens.

AIV could form an aerosol and horizontal transmission through the respiratory tract in poultry. In recent years, the dominant AIVs co-circulated in mainland China were H5 and H9 subtypes, of which H9 subtype AIVs have strong horizontal transmission. However, only some H5 AIVs had horizontal transmission ability (World Health Organization, 2013a). In our previous studies, some H5N1 AIVs, which belonged to clades 0, 2.3.2.2, 7.2, and 9, could horizontally transmit between chickens, ducks, geese, Japanese quails, and mice (Sun et al., 2011). Here, G2 and G3 virus belonged to clades 7.2 and 2.3.2.1, respectively, and both of them could transmit horizontally in chickens. All of the naive contact chickens in the G3 group had detected viruses shedding and replication in organs, but the naive contact chickens in the G2 group only had mild clinical symptoms and had no death or virus shedding. These results showed the new H5N1 HPAIVs of these two clades had different horizontal transmission ability. Moreover, H5N1 HPAIV of clade 2.3.2.1 still have been circulating in poultry and wild bird up to now, so they will continue to have the threat to human health and poultry product.

In China, poultry production modes, including rural household scatter breeding, poultry farms, and modern poultry ranches, are multiple and complicated so that prevention and control of H5 HPAI are difficult. Therefore, all poultry in China were immunized with H5N1 vaccinations compulsorily. In recent decades, several H5 vaccines, especially H5 inactivated vaccines, were widely used in China due to constant mutation and evolution of the virus (Swayne, 2012). The first commercial flu vaccine in China was an H5N2 inactivated vaccine, which used the low pathogenic avian influenza H5N2 virus A/Turkey/England/N-28/1973 and was approved for use in August of 2003 (Chen and Bu, 2009). Then the Re-1 vaccine was approved for use in 2004, for which was antigenically well-matched the epidemic strains at that time. In 2006, the H5N1 Re-4 vaccine, whose strain belonged to clade 7, was approved into service in China and widely used in the northern mainland. In 2008, new Re-5 vaccine began to be used in northern and southern China, whose strain A/duck/Anhui/1/2006(H5N1) belonged to clade 2.3.4 (Jiang et al., 2010; Li et al., 2010). In 2012, the recombinant vaccine Re-6 was also approved for use in the mainland to control new epidemic strains in clade 2.3.2. In 2013, a new H5 vaccine D7 was approved for use in waterfowl, which used an H5N2 virus (A/duck/Guangdong/D7/2007) belonged to clade 2.3.2. In conclusion, although flu vaccines were updated constantly, new strains still continue to appear in China. Therefore, it is necessary to evaluate the effectiveness and effects of current vaccines against the new strains timely.

From 2011 to 2012, most H5 isolates circulating in Guangdong province belonged to clade 2.3.2.1, including G3, so we wanted to estimate the effectiveness of current vaccines against them. In our study, the effectiveness of these three commercial vaccines against G3 varied. The D7 vaccine provided 100% protection to chickens against G3, the Re-6 vaccine provided 88.9% protection, and the Re-4 vaccine only provided 66.7%. The antibody titer of Re-4 in chicken had more 2 log2 than Re-6 and D7 when challenged at 28 DPV, but the protection rate of Re-4 against G3 was lowest because the Re-4 vaccine strain and G3 belonged to different clades. These results indicated that Re-4 vaccine did not protect chickens against H5 viruses challenging although had good immunogenicity and could induce high antibody levels. The D7 vaccine provided the best protection in these three vaccines against G3, whose strain belonged to clade 2.3.2. These told us that high antibody levels did not provide good protection, what need antigen matching between vaccine and epidemic strains. Therefore, to evaluate vaccines more objectively and effectively, we should be concerned not only about antibody level of immunized animals but also antigen matching between vaccine strains and epidemic isolates when observed the protection of vaccines in clinical practices.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: PJ, ML. Performed the experiments: XL, HS, YS, JC, NQ. Analyzed the data: PJ, HS, XL. Contributed reagents/materials/analysis tools: PJ, XL, YS, JC, JY, SW, NQ, TZ. Wrote the paper: HS, PJ. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (No. U1501212 and 31172343), the Program of International Science and Technology Cooperation of China (No. 2013DFA31940) the Science and Technology Projects of Guangdong Province (No. 2012B020306003,

2012A020100001 and 2014A050503061), the Science and Technology Projects of Guangzhou City (No. 201300000037 and 2013J4500030) and the Earmarked Fund for Modern Agro-Industry Technology Research System (CARS-42-G09).

REFERENCES

Chen, H., and Bu, Z. (2009). Development and application of avian influenza vaccines in China[J]. *Curr. Top. Microbiol. Immunol.* 333, 153–162. doi: 10.1007/978-3-540-92165-3_7

Chen, H., Deng, G., Li, Z., Tian, G., Li, Y., Jiao, P., et al. (2004). The evolution of H5N1 influenza viruses in ducks in southern China. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10452-10457. doi: 10.1073/pnas.0403212101

Claas, E. C., Osterhaus, A. D., van Beek, R., De Jong, J. C., Rimmelzwaan, G. F., Senne, D. A., et al. (1998). Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* 351, 472–477. doi: 10.1016/S0140-6736(97)11212-0

Fouchier, R. A., and Munster, V. J. (2009). Epidemiology of low pathogenic avian influenza viruses in wild birds[J]. *Rev Sci Tech*. 28, 49–58. doi: 10.20506/rst.28.1.1863

Gohrbandt, S., Veits, J., Hundt, J., Bogs, J., Breithaupt, A., Teifke, J. P., et al. (2011). Amino acids adjacent to the haemagglutinin cleavage site are relevant for virulence of avian influenza viruses of subtype H5. *J. Gen. Virol.* 92, 51–59. doi: 10.1099/vir.0.023887-0

Hoffmann, E., Stech, J., Guan, Y., Webster, R. G., and Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* 146, 2275–2289. doi: 10.1007/s007050170002

Jiang, W. M., Liu, S., Chen, J., Hou, G. Y., Li, J. P., Cao, Y. F., et al. (2010). Molecular epidemiological surveys of H5 subtype highly pathogenic avian influenza viruses in poultry in China during 2007-2009. *J. Gen. Virol.* 91, 2491–2496. doi: 10.1099/vir.0.023168-0

Jiao, P., Tian, G., Li, Y., Deng, G., Jiang, Y., Liu, C., et al. (2008). A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. *J. Virol.* 82, 1146–1154. doi: 10.1128/JVI.01698-07

Jiao, P., Wei, L., Song, Y., Cui, J., Song, H., Cao, L., et al. (2014). D701N mutation in the PB2 protein contributes to the pathogenicity of H5N1 avian influenza viruses but not transmissibility in guinea pigs. *Front. Microbiol.* 5:642. doi: 10.3389/fmicb.2014.00642

Lee, C. W., Suarez, D. L., Tumpey, T. M., Sung, H. W., Kwon, Y. K., Lee, Y. J., et al. (2005). Characterization of highly pathogenic H5N1 avian influenza A viruses isolated from South Korea. *J. Virol.* 79, 3692–3702. doi: 10.1128/JVI.79.6.3692-3702.2005

Li, Y., Shi, J., Zhong, G., Deng, G., Tian, G., Ge, J., et al. (2010). Continued evolution of H5N1 influenza viruses in wild birds, domestic poultry, and humans in China from 2004 to 2009. *J. Virol.* 84, 8389–8397. doi: 10.1128/JVI.00413-10

Nguyen, D. C., Uyeki, T. M., Jadhao, S., Maines, T., Shaw, M., Matsuoka, Y., et al. (2005). Isolation and characterization of avian influenza viruses, including highly pathogenic H5N1, from poultry in live bird markets in Hanoi, Vietnam, in 2001. *J. Virol.* 79, 4201–4212. doi: 10.1128/JVI.79.7.4201-4212.2005

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* 27, 493–497.

Reid, S. M., Shell, W. M., Barboi, G., Onita, I., Turcitu, M., Cioranu, R., et al. (2011). First reported incursion of highly pathogenic notifiable avian influenza A H5N1 viruses from clade 2.3.2 into European poultry. *Transbound. Emerg. Dis.* 58, 76–78. doi: 10.1111/j.1865-1682.2010.01175.x

Reperant, L. A., Rimmelzwaan, G. F., and Kuiken, T. (2009). Avian influenza viruses in mammals. *Rev. Sci. Tech.* 28, 137–159. doi: 10.20506/rst.28.1.1876

Sakoda, Y., Sugar, S., Batchluun, D., Erdene-Ochir, T. O., Okamatsu, M., Isoda, N., et al. (2010). Characterization of H5N1 highly pathogenic avian influenza virus strains isolated from migratory waterfowl in Mongolia on the way back from the southern Asia to their northern territory. *Virology* 406, 88–94. doi: 10.1016/j.virol.2010.07.007

Smith, G. J., Vijaykrishna, D., Ellis, T. M., Dyrting, K. C., Leung, Y. H., Bahl, J., et al. (2009). Characterization of avian influenza viruses A (H5N1) from wild birds, Hong Kong, 2004-2008. *Emerg. Infect. Dis.* 15, 402–407. doi: 10.3201/eid1503.081190

Sturm-Ramirez, K. M., Ellis, T., Bousfield, B., Bissett, L., Dyrting, K., Rehg, J. E., et al. (2004). Reemerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. *J. Virol.* 78, 4892–4901. doi: 10.1128/JVI.78.9.4892-4901.2004

Subbarao, K., Klimov, A., Katz, J., Regnery, H., Lim, W., Hall, H., et al. (1998). Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* 279, 393–396. doi: 10.1126/science.279.5349.393

Sun, H., Jiao, P., Jia, B., Xu, C., Wei, L., Shan, F., et al. (2011). Pathogenicity in quails and mice of H5N1 highly pathogenic avian influenza viruses isolated from ducks. *Vet. Microbiol.* 152, 258–265. doi: 10.1016/j.vetmic.2011.05.009

Swayne, D. E. (2012). The role of vaccines and vaccination in high pathogenicity avian influenza control and eradication[J]. *Expert Rev. Vaccines* 11, 877–880. doi: 10.1586/erv.12.60

Takatsy, G., and Barb, K. (1973). Antigenic analysis of prototype influenza a (H3N2) strains by the antiserum absorption method[J]. *J. Hyg. (Lond.)* 71, 501–508. doi: 10.1017/S0022172400046489

U.S. Geological Survey (2011). List of *Species Affected* by H5N1 (*Avian Influenza*). Reston, VA: U.S. Geological Survey—National Wildlife Health Center.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56, 152–179.

World Health Organization (2005). *Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1) Reported to WHO*. Geneva: WHO. Available online at: www.who.int/csr/disease/avian/influenza/country/cases/table/2005/06/28/en/index.html

World Health Organization/World Organization for Animal Health/Food and Agriculture Organization H5N1 Evolution Working Group (2008). Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1) (conference summary). *Emerg. Infect.*

World Health Organization (2012). *Antigenic and Genetic Characteristics of Zoonotic Influenza Viruses and Development of Candidate Vaccine Viruses for Pandemic Preparedness. 2011–2012.* Available online at: http://www.who.int/influenza/vaccines/virus/characteristics_virus_vaccines/en/.

World Health Organization (2013a). *Antigenic and Genetic Characteristics of Zoonotic Influenza Viruses and Development of Candidate Vaccine Viruses for Pandemic Preparedness. 2012-2013.* Available online at: http://www.who.int/influenza/vaccines/virus/characteristics_virus_vaccines/en/

World Health Organization (2013b). *Cumulative Number of Confirmed Human Cases for Avian Influenza A(H5N1) Reported to WHO, 2003-2013.* Available online at: http://www.who.int/influenza/human_animal_interface/EN_GIP_20130829CumulativeNumberH5N1cases

World Health Organization (2015). *Influenza at the Human-Animal Interface.* Summary and Assessment as of 1 May 2015. Available online at: http://www.who.int/influenza/human_animal_interface/Influenza

Xu, X., Subbarao, Cox, N.J., and Guo, Y. (1999). Genetic characterization of the pathogenic influenza A/Goose/Guangdong/1/96 (H5N1) virus: similarity of its hemagglutinin gene to those of H5N1 viruses from the 1997 outbreaks in Hong Kong. *Virology* 261, 15–19. doi: 10.1006/viro.1999.9820

Yuan, R., Cui, J., Zhang, S., Cao, L., Liu, X., Kang, Y., et al. (2014). Pathogenicity and transmission of H5N1 avian influenza viruses in different birds. *Vet. Microbiol.* 168, 50–59. doi: 10.1016/j.vetmic.2013.10.013

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Jiao, Song, Liu, Song, Cui, Wu, Ye, Qu, Zhang and Liao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 20 June 2017
doi: 10.3389/fmicb.2017.01081

Immune Responses of Chickens Infected with Wild Bird-Origin H5N6 Avian Influenza Virus

Shimin Gao[1,2†], Yinfeng Kang[2,3†], Runyu Yuan[2,4†], Haili Ma[1], Bin Xiang[2], Zhaoxiong Wang[5], Xu Dai[2], Fumin Wang[6], Jiajie Xiao[6], Ming Liao[2] and Tao Ren[2]**

[1] College of Animal Science and Veterinary Medicine, Shanxi Agriculture University, Taigu, China, [2] College of Veterinary Medicine, Key Laboratory of Zoonosis Prevention and Control of Guangdong Province, South China Agricultural University, Guangzhou, China, [3] State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Department of Experimental Research, Sun Yat-sen University Cancer Center, Guangzhou, China, [4] Key Laboratory for Repository and Application of Pathogenic Microbiology, Research Center for Pathogens Detection Technology of Emerging Infectious Diseases, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China, [5] College of Animal Science, Yangtze University, Jingzhou, China, [6] Guangdong Provincial Wildlife Rescue Center, Guangzhou, China

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Xiaoming Sun,
Ragon Institute of MGH, MIT and Harvard, United States
Bin Su,
Center for Infectious Diseases, Beijing You'an Hospital, Capital Medical University, China
Ding Yuan Oh,
WHO Collaborating Centre for Reference and Research on Influenza (VIDRL), Australia

***Correspondence:**
Ming Liao
mliao@scau.edu.cn
Tao Ren
rentao6868@126.com

[†] These authors have contributed equally to this work and are co-first authors.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 05 April 2017
Accepted: 29 May 2017
Published: 20 June 2017

Citation:
Gao S, Kang Y, Yuan R, Ma H, Xiang B, Wang Z, Dai X, Wang F, Xiao J, Liao M and Ren T (2017) Immune Responses of Chickens Infected with Wild Bird-Origin H5N6 Avian Influenza Virus. Front. Microbiol. 8:1081. doi: 10.3389/fmicb.2017.01081

Since April 2014, new infections of H5N6 avian influenza virus (AIV) in humans and domestic poultry have caused considerable economic losses in the poultry industry and posed an enormous threat to human health worldwide. In previous research using gene sequence and phylogenetic analysis, we reported that H5N6 AIV isolated in February 2015 (ZH283) in Pallas's sandgrouse was highly similar to that isolated in a human in December 2015 (A/Guangdong/ZQ874/2015), whereas a virus (i.e., SW8) isolated in oriental magpie-robin in 2014 was highly similar to that of A/chicken/Dongguan/2690/2013 (H5N6). However, the pathogenicity, transmissibility, and host immune-related response of chickens infected by those wild bird-origin H5N6 AIVs remain unknown. In response, we examined the viral distribution and mRNA expression profiles of immune-related genes in chickens infected with both viruses. Results showed that the H5N6 AIVs were highly pathogenic to chickens and caused not only systemic infection in multiple tissues, but also 100% mortality within 3–5 days post-infection. Additionally, ZH283 efficiently replicated in all tested tissues and transmitted among chickens more rapidly than SW8. Moreover, quantitative real-time polymerase chain reaction analysis showed that following infection with H5N6, AIVs immune-related genes remained active in a tissue-dependent manner, as well as that ZH283 induced mRNA expression profiles such as *TLR3*, *TLR7*, *IL-6*, *TNF-α*, *IL-1β*, *IL-10*, *IL-8*, and *MHC-II* to a greater extent than SW8 in the tested tissues of infected chickens. Altogether, our findings help to illuminate the pathogenesis and immunologic mechanisms of H5N6 AIVs in chickens.

Keywords: influenza virus, H5N6, wild birds, pathogenicity, transmissibility, immune response

INTRODUCTION

Since 2003, multiple highly pathogenic avian influenza A (HPAI) H5 subtypes, including H5N1, H5N2, H5N6, and H5N8, have generated severe epidemics and thus not only tremendous economic losses in the domestic poultry industry, but also serious threats to human health worldwide (Jhung and Nelson, 2015). As of October 3, 2016, at least 856 cases of human infection

with avian influenza A (H5N1) virus in 16 countries had been reported to the World Health Organization, among which 452 had ended in death, for an apparent case fatality rate of 52.8% (WHO, 2016). As the natural reservoir for avian influenza viruses (AIVs), wild bird populations can be infected by many such viruses, including the H3, H5, and H7 subtypes AIVs, and thus play a critical role in AIV epidemiology and ecology (Claes et al., 2016; Dhingra et al., 2016; Kang et al., 2017). Thus far, results of the phylogenetic analysis of the hemagglutinin (HA) gene have revealed multiple clades and subclades of H5 subtype AIVs. Among them, H5N6 has replaced H5N1 as the dominant subtype in southern China (Bi et al., 2016a), while clade 2.3.4.4 of AIVs is now considered to be the dominant in China (Saito et al., 2015; Claes et al., 2016). Given recent suggestions that clade 2.3.4.4 of AIVs has become increasingly pathogenic to domestic poultry and wild birds (Claes et al., 2016; Sun et al., 2016), AIV virulence is likely affects multiple factors and depends upon both antigenic drift and the AIV-infected strain in the host immunity (Tscherne and Garcia-Sastre, 2011).

An AIV replicates primarily in the respiratory system (Sturm-Ramirez et al., 2004), from where it spreads to the brain and lymphoid tissues by way of infection. Such infection induces batteries of receptors and triggers a signaling cascade that ultimately activates the host's immune response. As part of that process, for example, the endosomal Toll-like receptor (TLR) 3 and sphingosine-1-phosphate-1 receptor (S1PR1) recognize double-stranded viral RNA released during the uncoating of an internalized virus (Barton, 2007; Teijaro et al., 2011). During AIV infection in mammals, the endosomal TLR 7/8, which recognizes single-stranded viral RNA, can prompt the production of interferon (IFN)-α and IFN-β (Diebold et al., 2004). As MacDonald et al. (2008) have shown, when TLR7/8 are activated by AIV infection in host cells, the recognition of viral RNA results in the secretion of proinflammatory cytokines (e.g., IL-1β, IL-6) and antiviral cytokines (e.g., IFNs). By extension, the expression of proinflammatory cytokines and IFNs influences both viral clearance and the manifestation of clinical symptoms. At the same time, since major histocompatibility complex (MHC) classes I and II antigen presentation molecules used for AIV uptake activated cellular immunity and humoral immunity of B cells (i.e., IFN-γ) and T cells (i.e., CD3+, CD4+, CD8+), MHC molecules likely play a role in activating host innate immune response to AIV infection (Gromme and Neefjes, 2002; Williams et al., 2002).

In China's Sichuan Province on May 7, 2014, the first-ever fatal case of human infection by a reassortant H5N6 AIV involved a 49-year-old man with a history of exposure to live poultry. To date, 14 additional cases of human infection with the H5N6 virus in China's Sichuan, Guangdong, Jiangxi, and Yunnan Provinces—10 of which ended in death—documented by the World Health Organization and World Organisation for Animal Health were characterized as posing a potential risk to public health[1].

In studies conducted during 2014–2015, we performed epidemic surveillance of AIVs among wild birds at nature reserves in southern China, isolated two novel reassortant HPAI H5N6 viruses, and conducted genetic and phylogenetic analyses to elucidate their molecular features (Kang et al., 2017). By extension, in the present study, we investigated the pathogenicity and transmissibility of the viruses in chickens. In addition, to assess the role of the host innate immune response of H5N6-infected chickens, we examined a complex expression profile of pattern recognition receptors (PRRs), proinflammatory cytokines, chemokines, and MHC molecules in the brain, lung, spleen, and bursa of Fabricius.

[1]http://www.who.int/en/

MATERIALS AND METHODS

Ethics Statement

All animal experiments were conducted in ABSL-3 facilities and in accordance with the guidelines of South China Agricultural University's Institutional Animal Care and Use Committee. All animal protocols were approved by the Committee on the Ethics of Animal Experiments of the ABSL-3 Committee of South China Agricultural University (approval no. L102012017001K).

Viruses and Experimental Animals

Two H5N6 viruses—namely, A/oriental magpie-robin/Guangdong/SW8/2014 (SW8) and A/Pallas's sandgrouse/Guangdong/ZH283/2015 (ZH283)—used in this study were grown and purified three times in Madin–Darby canine kidney cells by standard plaque assay. The stocks of H5N6 viruses were propagated in 9-day-old specific pathogen-free (SPF) chicken eggs at 37°C for 72 h per the procedure (Yuan et al., 2014). Allantoic fluid pooled from multiple eggs was taken for centrifugation for 2 min at 8,000 rpm, from which the supernatant was harvested and subsequently frozen in aliquots at −80°C for further characterization. The 50% egg infectious dose (EID_{50}) titer for egg-grown virus was determined by 10-fold serial dilutions and the titration of each virus in 9-day-old SPF eggs using Reed and Muench's (1938) method. Six-week-old SPF white leghorn chickens (Guangdong Wens Dahuanong Biotechnology Co., Ltd, Yunfu, China) were held in isolator cages with a feeding space of 117 m^3 throughout the duration of each experiment.

Pathogenesis and Transmission Experiments of H5N6 Virus in Chickens

In vivo pathogenesis studies of wild bird H5N6 influenza viruses were designed as previously described (Zhang et al., 2008, 2009; Pu et al., 2015). In brief, groups of 12 6-week-old SPF chickens were intranasally inoculated with 0.2 mL of 10^5 EID_{50} of SW8 or ZH283, while a control group of 12 chickens was inoculated with 0.2 mL of phosphate buffered saline (PBS) using the same route. Three days later, six inoculated chickens from each group were humanely euthanized to test for viral replication in lung, kidney, spleen, cecal tonsils, bursa of Fabricius, trachea, pancreas, liver, heart, brain, duodenum, ileum, descending colon, and jejunum tissue. The remaining chickens were observed twice daily, at 8:00 and 20:00, for clinical symptoms, morbidity, and mortality

for 14 days according to the protocol provided by the World Organisation for Animal Health[2].

Direct contact virus transmission experiments in chickens were conducted per the procedure of (Yuan et al., 2014). Briefly, the chickens of inoculated groups ($n = 6$) were intranasally inoculated with 0.2 mL of 10^5 EID_{50} of either the SW8 or ZH283 virus in a ABSL-3 laboratory, and after 24 h, additional naïve contact groups ($n = 3$) were also intranasally inoculated with 0.2 mL of PBS and placed in physical contact in the same cage to share feed and water with chickens inoculated with the virus. At 3 days post-infection (DPI), three inoculated chickens were humanely euthanized, and target tissues (i.e., brain, lung, spleen, and bursa of Fabricius) were harvested to determine viral titers and for RNA extraction. At 3, 5, 7, 9, and 11 DPI, oropharynx and cloacal swabs samples were collected for the detection of viral shedding and suspended in 1 mL of PBS. All tested tissues and swabs samples were harvested for viral detection and titration in SPF chick embryos. All surviving chickens were euthanized at 14 DPI, and the serum was harvested and tested for seroconversion by hemagglutination inhibition testing using 1% turkey erythrocytes (Stephenson et al., 2004).

RNA and cDNA Preparation

Total RNA was extracted from the brain, lung, spleen, and bursa of Fabricius of H5N6-inoculated chickens and mock-infected chickens at 3 DPI using the Takara MiniBEST Universal RNA Extraction Kit (Takara Bio Inc., Tokyo, Japan) following the manufacturer's instructions. Total RNA (1 μg) was reverse-transcribed with the PrimeScript™ II 1st Strand cDNA Synthesis Kit (Takara Bio Inc.) and stored at −20°C for further study.

Quantitative Real-Time Polymerase Chain Reaction

Quantitative real-time polymerase chain reaction (qRT-PCR) was performed using a FastStart Universal SYBR Green Master kit (Roche Diagnostics, Shanghai, China). qRT-PCR primers (**Table 1**) were designed from published target sequences and previously reported (Adams et al., 2009) with Primer Premier 7.0 software (Premier Biosoft, Palo Alto, CA, United States). qRT-PCR was performed on a LightCycler480 (Roche Applied Science, Mannheim, Germany), the products of which were purified by using a DNA gel extraction kit (Takara Bio Inc., Tokyo, Japan). For the purposes of assay validation, purified products were cloned into pMD19-T and sequenced to verify correct target amplification.

Calculations and Statistical Analysis

The relative expression ratios of target genes in tested tissues vs. those in control tissues were calculated by the $2^{-\Delta\Delta CT}$ method using the chicken housekeeping gene glyceraldehyde-3-phosphate-dehydrogenase (NM_204305) as the endogenous reference gene in order to normalize the level of target gene expression (Livak and Schmittgen, 2001). Standard deviations were determined by using the relative expression ratios of three replicates for each gene measured. Differences of virus titers and mRNA expression levels were statistically analyzed with an unpaired non-parametric test and paired Student's *t*-test, respectively, using GraphPad Prism version 6.0 (GraphPad Software Inc., La Jolla, CA, United States) software. Compared to the mock-infected control, $p < 0.05$ and $p < 0.01$ were considered to indicate a statistically significant difference unless stated otherwise.

RESULTS

Pathogenesis of Wild Birds Origin A(H5N6) Influenza Viruses in Chickens

In previous research, we characterized the two H5N6 influenza viruses isolated from apparently healthy wild birds in 2014 and 2015 in Guangdong Province, China. On the one hand, SW8 was isolated from an oriental magpie-robin, and its PB2 gene with poultry H5N6 viruses shared the highest nucleotide similarity with that of A/chicken/Dongguan/2690/2013 (H5N6). On the other hand, ZH283 was isolated from a Pallas's sandgrouse, and its PB2 gene shared the highest nucleotide similarity with that of A/Guangdong/ZQ874/2015 (H5N6) isolated from a 40-year-old woman who reported exposure to domestic poultry[3].

To determine the pathogenicity of the viruses in chickens, we intranasally inoculated 6-week-old SPF white leghorn chickens with 10^5 EID_{50} of either H5N6 virus (i.e., SW8 or ZH283). All inoculated chickens exhibited clinical signs of illness, including severe depression, cloudy eye, and intermittent head-shaking, and died within 5 DPI, with a mean death time (MDT) of 3.3 to 4.0 d (**Figure 1A**). SW8 and ZH283 replicated systemically in chickens and at 3 DPI was detectable in all tested organs, including the respiratory tract (i.e., lung and trachea), kidney, lymphoid tissues (i.e., spleen, cecal tonsils, and bursa of Fabricius), pancreas, liver, brain, intestinal tract (i.e., duodenum, ileum, descending colon, and jejunum), and heart. SW8 and ZH283 replicated efficiently in the lower respiratory tract; high viral titers were detected in the lung, with mean titers of 6.33 $\log_{10}EID_{50}$ and 8.58 $\log_{10}EID_{50}$, respectively (**Figure 1B**). The two novel viruses also replicated in the brain, spleen, and bursa of Fabricius, with mean titers of 4.83–7.17 $\log_{10}EID_{50}$, 5.83–7.33 $\log_{10}EID_{50}$, and 6.08–7.58 $\log_{10}EID_{50}$, respectively (**Figure 1B**). Overall, the two novel H5N6 influenza viruses of wild bird origin showed high pathogenicity in chickens and could replicate systematically in them.

Transmissibility of A(H5N6) Influenza Viruses of Wild Bird Origin in Chickens

To evaluate the horizontal intraspecies transmissibility of the two novel H5N6 viruses, three SPF chickens were intranasally inoculated with 0.2 mL PBS and introduced into the same cage

[2]http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.03.04_AI.pdf

[3]http://www.who.int/csr/don/4-january-2016-avian-influenza-china/en/

TABLE 1 | Quantitative real-time PCR primers used in this study.

Gene	Forward primer (5′3′)	Reverse primer (5′3′)	Product size (bp)	GenBank accession no.
GAPDH	CCTCTCTGGCAAAGTCCAAG	CATCTGCCCATTTGATGTTG	200	NM_204305
TLR3	ACAATGGCAGATTGTAGTCACCT	GCACAATCCTGGTTTCAGTTTAG	123	NM_001011691
TLR7	TGTGATGTGGAAGCCTTTGA	ATTATCTTTGGGCCCCAGTC	218	DQ780342
S1PR1	TTGCCTTGTCAGCTTCTGTG	CGTGGAGCAGTTTGACAAGA	203	XM_422305
IL-1β	GCTCTACATGTCGTGTGTGATGAG	TGTCGATGTCCCGCATGA	80	NM_204524
IL-6	CCTGTTCGCCTTTCAGACCT	GGGATGACCACTTCATCGGG	171	EU170468
IL-8	ATTCAAGATGTGAAGCTGAC	AGGATCTGCAATTAACATGAGG	196	DQ393272
TNF-α	CCGCCCAGTTCAGATGAGTT	GCAACAACCAGCTATGCACC	130	AY765397
IFN-α	ATGCCACCTTCTCTCACGAC	AGGCGCTGTAATCGTTGTCT	387	EU367971
IFN-β	CCTCAACCAGATCCAGCATT	GGATGAGGCTGTGAGAGGAG	259	AY831397
IFN-γ	TGAGCCAGATTGTTTCGATG	CTTGGCCAGGTCCATGATA	248	DQ906156
CCL5	GTTTGGGGCTGATACAACCG	CCTTCACATGATTCTGGGGCA	71	NM_001045832
MHC-I	AAGAAGGGGAAGGGCTACAA	AAGCAGTGCAGGCAAAGAAT	222	NM_001031338
MHC-II	CTCGAGGTCATGATCAGCAA	TGTAAACGTCTCCCCTTTGG	312	DQ008588

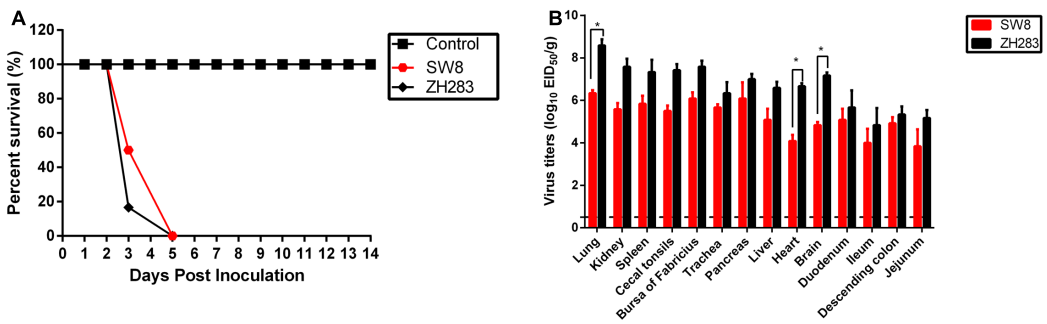


FIGURE 1 | Pathogenesis of H5N6 viruses in specific pathogen-free chickens. **(A)** Percentage of survival of SW8 and ZH283 in chickens. **(B)** Comparison of two A(H5N6) influenza virus titers of wild bird origin in chickens. Groups of 12 6-week-old chickens were intranasally inoculated with 0.2 mL of 10^5 EID_{50} of SW8, ZH283, or PBS; six chickens in each group were euthanized at 3 days post-infection, and lung, kidney, spleen, cecal tonsils, bursa of Fabricius, trachea, pancreas, liver, heart, brain, duodenum, ileum, descending colon, and jejunum tissues were collected. The remaining chickens were observed for clinical signs of illness and lethality for 2 weeks. Virus titers were determined in eggs and expressed as $\log_{10}$ EID_{50}/g of tissue. Data are expressed as $M \pm SD$. Dashed black lines indicate the lower limit of detection. Differences were analyzed with a paired Student's *t*-test and were considered statistically significant at $^*p < 0.05$ compared to control.

as a naïve contact group, which were then housed with chickens inoculated with SW8 or ZH283. Shedding of SW8 could be detected from both oropharyngeal and cloacal swabs within 3 DPI, with viral titers in the ranges of 2.42–3.83 $\log_{10}EID_{50}$ in oropharyngeal swab samples and of 1.52–3.79 $\log_{10}EID_{50}$ in cloacal swab samples. ZH283 could also be detected from oropharyngeal and cloacal swabs within 5 DPI, with viral titers in the range of 4.58–4.75 $\log_{10}EID_{50}$ in oropharyngeal swabs and of 3.50–3.90 $\log_{10}EID_{50}$ in cloacal swabs (**Figures 2A,B**). Naïve contact chickens co-housed with chickens inoculated with SW8 did not die during the observation time, but all contact group chickens seroconverted and exhibited high titers (9.33 ± 0.58 log2), as shown in **Table 2**. Viral shedding was observed in both oropharyngeal and cloacal swabs, and viral titers of 1.50–1.83 $\log_{10}EID_{50}$ within 5 DPI were detected in oropharyngeal swabs (**Figure 2A**); however, viral titers of the cloacal swabs could be detected (1.08 $\log_{10}EID_{50}$) at 3 DPI (**Figure 2B**). Naïve contact chickens co-housed with chickens inoculated with ZH283 exhibited 100% lethality and mortality, with a MDT of 5.0 days (**Table 2**), and exhibited clinical signs of illness, including coughing, cloudy eye, and dyspnea. All surviving chickens in the naïve contact group co-housed with ZH283-infected chickens shed virus from the oropharynx and cloaca within 7 DPI, with mean viral titers of 2.75–3.75 $\log_{10}EID_{50}$ in oropharyngeal swabs and of 1.75–4.50 $\log_{10}EID_{50}$ in cloacal swabs (**Figures 2A,B**). In short, results demonstrate that the two novel H5N6 influenza viruses replicated efficiently in chickens and exhibited efficient transmission via direct contact in the chicken model.

Expression of TLRs and S1PR1 in the Target Tissues of H5N6-Infected Chickens

Toll-like receptors are PRRs with a unique and essential physiological function in host immune systems activated by pathogen-associated molecular patterns (Medzhitov, 2001). Expression profiles of two PRRs—*TLR3* and *TLR7*—were examined in the target tissues of H5N6-infected chickens. As shown in **Figure 3A**, in contrast to mock-infected chickens,

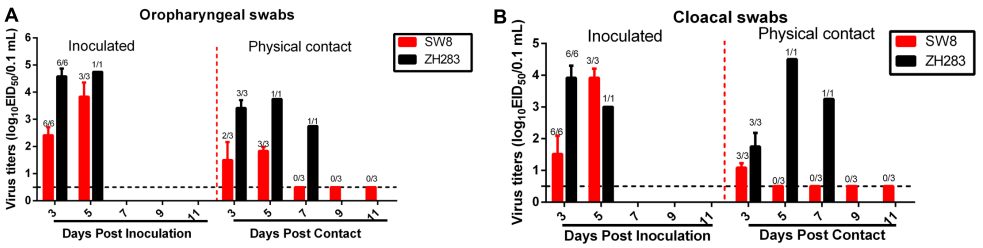


FIGURE 2 | Direct contact transmissibility of H5N6 influenza viruses of wild bird origin among chickens. Viral titers of ZH283 and SW8 in oropharyngeal swabs **(A)** and cloacal swabs **(B)** in H5N6 influenza virus-inoculated and physical contact chickens. Three chickens were inoculated intranasally with 10^5 EID_{50} of SW8 or ZH283, whereas three naïve chickens were placed in the cage of H5N6-infected chickens at 24 h post-infection to initiate contact. Oropharyngeal and cloacal swabs were collected from infected and naïve contact chickens at indicated time points; virus titers were titrated and are expressed as $log_{10}EID_{50}$/0.1 mL. Data are expressed as $M \pm SD$. The proportion of chicken swabs presenting infectious virus from all detected swabs at indicated time points appears in the figure above each group. Dashed black lines indicate the lower limit of detection.

TABLE 2 | Illness, mortality and HI titers of SPF chickens response to H5N6 influenza virus infection[a].

Strains	Titer ($log_{10}EID_{50}$)	Group	Illness[b]	Mortality (%)	HI titer[c] (log2, mean ± SD)	MDT
A/oriental magpie-robin/Guangdong/ SW8/2014 (H5N6)	7.88	Inoculated	3/3	3/3 (100)	–[d]	4.0
		Contact[e]	0/3	0/3	3/3 (9.33 ± 0.58)	
A/Pallas's sandgrouse/ Guangdong/ZH283/2015 (H5N6)	8.50	Inoculated	3/3	3/3 (100)	–	3.3
		Contact	3/3	3/3 (100)	–	5.0
Controls (no virus exposure)			0/3	0/3	0/3	

[a]*Unless indicated otherwise, data represent the number of affected animals/animals in the group. Six-week-old chickens were inoculated by the intranasal route with 10^5 EID_{50} of each virus in a 0.2 mL volume; HI, hemagglutination inhibition; MDT, mean death time.*
[b]*Severe depression, coughing, cloudy eye, dyspnea and intermittent head-shaking.*
[c]*HI titer was assayed in serum samples taken at 14 days post-inoculation. Data show the ratio of antibody-positive chickens to the number of virus-inoculated chickens.*
[d]*All the chickens died at the end of the observation.*
[e]*Three additional naïve contact chickens were placed with inoculated chickens as a contact group 24 h after inoculation.*

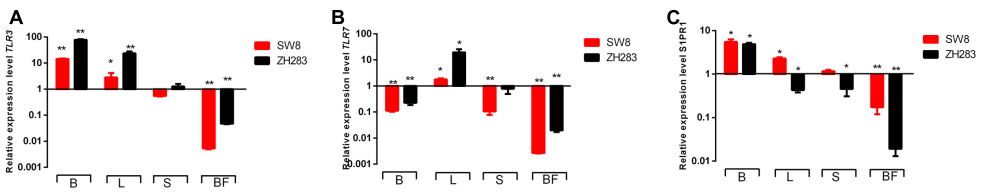


FIGURE 3 | Toll-like receptors (TLRs) and sphingosine-1-phosphate-1 receptor (S1RP1) expression profiles in the target tissues of chickens infected with H5N6. At 3 days post-infection, the target tissues (i.e., brain, lung, spleen, and bursa of Fabricius) of H5N6-infected chickens were harvested for TLR and S1RP1 mRNA level detection via qRT-PCR method. **(A)** *TLR3*, **(B)** *TLR7*, **(C)** *S1PR1*. Data are expressed as $M \pm SD$. Differences were analyzed with a paired Student's *t*-test and were considered statistically significant at $^*p < 0.05$, $^{**}p < 0.01$ compared to control. B, Brain; L, Lung; S, Spleen; BF, Bursa of Fabricius.

their expression level of *TLR3* in the brain and lung was significantly elevated when induced by both viruses, with a fold increase of 2.80–78.56 in the brain or lung. In the spleen, the expression level of *TLR3* was downregulated in response to SW8 infection, yet upregulated following infection with ZH283. In the bursa of Fabricius, the expression level of *TLR3* was markedly downregulated when induced by both viruses. The expression level of *TLR7* was upregulated in the lung when induced by SW8 or ZH283, by 1.79- and 19.41-fold, respectively. However, the expression level of *TLR7* in the brain, spleen, and bursa of Fabricius showed different expression patterns when induced by the viruses; *TLR7* expression level was downregulated when induced by both viruses compared to the control, with a fold change of 0.003–0.78 in all tested tissues except lung tissue. In particular, *TLR7* expression remained low and was no longer visible in the bursa of Fabricius when triggered by both viruses. Notably, the expression levels of *TLR3* and *TLR7* in target tissues induced by ZH283 were generally greater than those induced by SW8 (**Figures 3A,B**).

As an indispensable regulator of inflammation activation, S1PR1 plays a crucial role in immune cell trafficking and immune response (Rivera et al., 2008). When induced by SW8, S1PR1 expression was upregulated in the brain, lung, and spleen—by 5.51-, 2.29-, and 1.16-fold, respectively—but not in the bursa of Fabricius (0.17-fold). However, the expression level S1PR1 showed different tendencies when infected by ZH283. Unlike the expression level of TLR3 and TLR7, S1PR1 expression in the tested tissues after infection with ZH283 was lower than that in

response to infection with SW8 (**Figure 3C**). Our data indicate that the engagement of PRRs and S1PR1 by the H5N6 influenza virus occurs in a tissue-dependent manner.

Expression of Proinflammatory Cytokines and Chemokines in the Target Tissues of H5N6-Infected Chickens

The engagement of TLRs by influenza virus in specific target tissues initiated animal immunity via the production of proinflammatory cytokines and chemokines, including *IL-1*β, *IL-6*, *IL-8*, *IL-10*, *TNF*-α, and *CCL5*. As shown in **Figures 4A,B,D**, the expression level of *IL-1*β, *IL-6*, and *IL-10* were remarkably unregulated in the lungs of tested chickens when infected by SW8 and ZH283 compared to those of mock-infected chickens. On the contrary, in the brain, spleen, and bursa of Fabricius, the expression levels of *IL-1*β, *IL-6*, and *IL-10* were downregulated when induced by both viruses. However, the expression levels of *IL-8*, *TNF*-α, and *CCL5* in the tested tissues of infected chickens showed a different expression patterns. As illustrated in **Figure 4F**, ZH283 induced an upregulated expression level of *CCL5* in all tested tissues, whereas SW8 induced an upregulated expression level of *CCL5* in the brain and spleen, but a downregulated one in the lung and bursa of Fabricius. Notably, ZH283-induced expression levels of *IL-1*β, *IL-8*, *TNF*-α, *IL-6*, and *IL-10* were greater than those induced by SW8 in all tested tissues of chickens (**Figures 4A–E**).

The activation of TLRs also mediated the activation of IFN regulatory factor 3/7, primarily by recruiting MyD88 or TNF receptor-associated factor 6, which ultimately activated I and II IFNs (i.e., *IFN*-α, *IFN*-β, and *IFN*-γ). In the lungs of tested chickens, both ZH283 and SW8 induced significantly upregulated expression levels of *IFN*-α, *IFN*-β, and *IFN*-γ by 7.55- and 75.97-fold, 68.23- and 362.80-fold, 30.11- and 85.31-fold, respectively ($p < 0.05$) compared to uninoculated chickens (**Figures 4G–I**). In contrast to the lung, the brain, spleen, and bursa of Fabricius showed different expression patterns in the levels of *IFN*-α, *IFN*-β, and *IFN*-γ in response to ZH283 and SW8 infection. However, ZH283 induced the expression levels of *IFN*-α, *IFN*-β, and *IFN*-γ to a greater extent than SW8 in the tested tissues of infected chickens.

In sum, our data indicate that the mRNA expression profiles of proinflammatory cytokines and chemokines showed different patterns in tested tissues likely associated with the pathogenic difference of both viruses in chickens.

Expression of MHC Classes I and II Molecules in the Target Tissues of H5N6-Infected Chickens

To investigate whether MHC classes I and II molecules were involved in the host innate immune response to H5N6 influenza virus infection, we examined their expression levels in the lung, brain, spleen, and bursa of Fabricius in chickens at 3 DPI. As

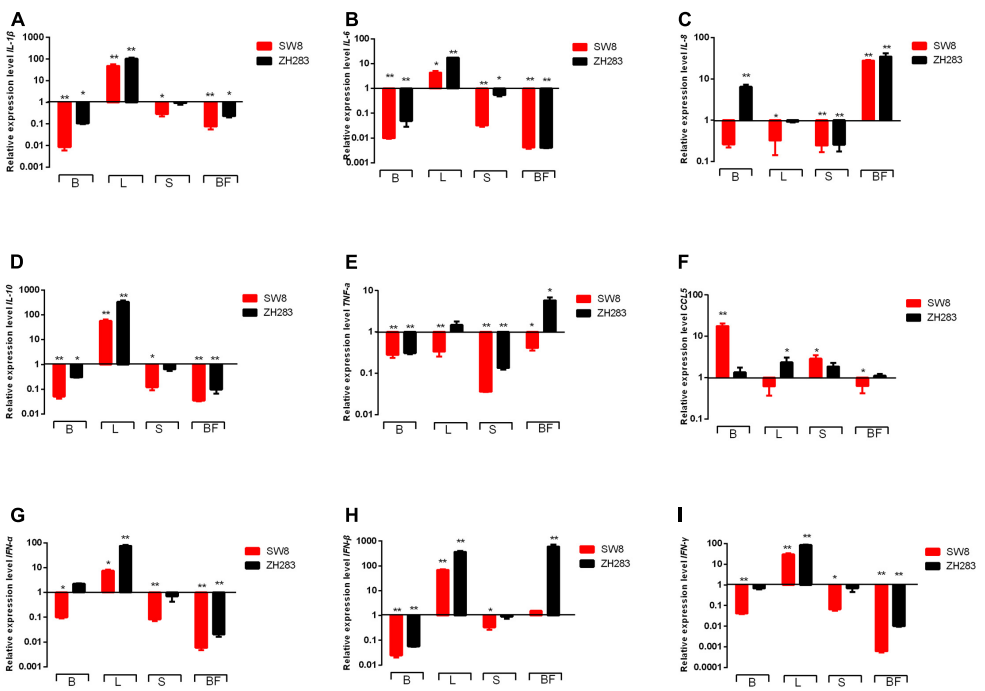


FIGURE 4 | Proinflammatory cytokines and chemokines expression profiles in the target tissues of chickens when infected by H5N6. At 3 days post-infection, the target tissues (i.e., brain, lung, spleen, and bursa of Fabricius) of H5N6-infected chickens were harvested for proinflammatory cytokines and chemokines mRNA level detection via quantitative real-time polymerase chain reaction. **(A)** *IL-1*β, **(B)** *IL-6*, **(C)** *IL-8*, **(D)** *IL-10*, **(E)** *TNF*-α, **(F)** *CCL5*, **(G)** *IFN*-α, **(H)** *IFN*-β, **(I)** *IFN*-γ. Data are expressed as $M \pm SD$. Differences were analyzed with a paired Student's *t*-test and considered statistically significant at $^{*}p < 0.05$, $^{**}p < 0.01$ compared to the control. B, Brain; L, Lung; S, Spleen; BF, Bursa of Fabricius.

illustrated in **Figures 5A,B**, MHC classes I and II molecule expression levels were upregulated in the brain, spleen, and bursa of Fabricius when infected by both viruses. In the lung, in contrast to the mock-infected control, the expression level of the MHC class I molecule was remarkably downregulated (0.063- and 0.20-fold, respectively, $p < 0.05$); however, that of the MHC class II molecule was significantly upregulated when induced by SW8 and ZH283 (12.83- and 99.08-fold, respectively, $p < 0.05$). Those results demonstrated that MHC classes I and II molecules could play a significant role in the course of host innate immune response to H5N6 influenza virus infection in chickens.

DISCUSSION

The first case of human infection with H5N6 AIVs was reported in southwest China's Sichuan Province in 2013 (Pan et al., 2016). Results of epidemiological surveillance show that the viruses have recently been isolated from humans (Shen et al., 2016), domestic poultry (Bi et al., 2015; Butler et al., 2016; Du et al., 2017; Li et al., 2017), pigs (Li et al., 2015), environmental samples (Yuan et al., 2016), cats (Yu et al., 2015), and wild birds (Bi et al., 2016b) and resulted in heavy losses in the poultry industry. However, the pathogenicity and transmissibility of H5N6 AIVs have remained unclear. In the current research, we systematically investigated the pathogenicity, transmissibility and the host immune-related gene in the target tissues of infected chickens when challenged by those of wild bird-origin H5N6 AIVs. Our findings provide insights into understanding the host innate immune response of chickens to infection with different pathogenicities of wild bird-origin H5N6 AIVs.

Importantly, we found that both H5N6 viruses isolated from wild birds were highly pathogenic and could efficiently be transmitted in chickens. Both viruses were shed from the oropharynx and cloaca in inoculated chickens and could be efficiently transmitted from infected chickens to naïve contact groups, the latter of which also shed viruses from both the cloaca and oropharynx throughout the experimental period. That the H5N6 HPAIVs isolated from wild birds could infect and be transmitted in chickens suggests that they may co-circulate in poultry and thus pose a great threat to the poultry industry.

Notably, chickens inoculated with SW8 showed high pathogenicity, whereas naïve contact chickens infected showed no deaths. By contrast, chickens inoculated with ZH283 showed high pathogenicity, with a mortality rate of 100% within 2–3 days and efficient horizontal transmission in chickens. The mechanisms of lethality and transmissibility might be associated with mutations at positions K52T, I155T, and A544V of the HA protein, at positions K207R and Y436H of the PB1 protein, and at position T515A of the PA protein (Hulse-Post et al., 2007; Li et al., 2014). However, with the exception of position I155T of the HA protein, no mutations were observed in ZH283, which suggests that differences in the pathogenicity and transmissibility of H5N6 influenza viruses in chickens correlate with the probability of their being at position I155T of the HA protein. In addition, the transmissibility of H5N6 AIV in different birds may also depend on the stability of viral particles and the difference of viral protein structure, relative humidity, and temperature (Webster et al., 1992; Lowen et al., 2007). However, our experiment posed several limitations, meaning that more viral strains isolated from different animals and species need to be tested in order to investigate the correlation between pathogenicity and host immunity. Further investigation is also clearly needed to elucidate the differences of pathogenicity, transmissibility, and host innate immune response to infection with H5N6 AIVs in chickens.

Remarkably, the expression levels of *TLR3* and *S1PR1* were upregulated in the brain following infection with SW8 and ZH283, yet showed different expression patterns in lymphoid tissues. Similarly, the production of *TLR7*, *IL-1*β, *IL-6*, *IL-10*, and *IFN-*γ were upregulated in the lung but downregulated in brain, spleen, and bursa of Fabricius in response to both viruses. Such results suggest that the engagement of the TLRs and cytokines are involved in a tissue-dependent manner. Previous studies have revealed tissue-specific immune responses following infection with H5N1 (Wei et al., 2013), H5N2 (Vanderven et al., 2012), and H7N1 (Cornelissen et al., 2012). The difference of cell types could be associated with immune responses and virus titers in the tissues tested for infection.

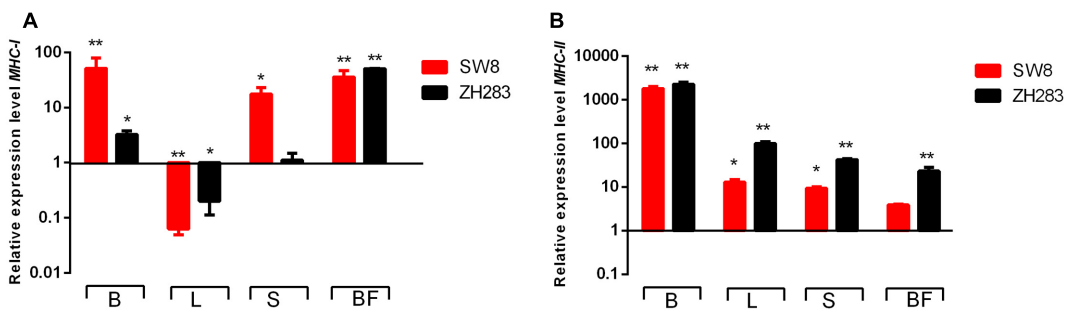


FIGURE 5 | Major histocompatibility complex (MHC) classes I and II molecule expression profiles in the target tissues of chickens infected by H5N6. At 3 days post-infection, target tissues (i.e., brain, lung, spleen, and bursa of Fabricius) of H5N6-infected chickens were harvested for MHC classes I and II molecule mRNA level detection via quantitative real-time polymerase chain reaction. **(A)** *MHC-I*, **(B)** *MHC-II*. Data are expressed as $M \pm SD$. Differences were analyzed with a paired Student's *t*-test and considered statistically significant at $^{*}p < 0.05$, $^{**}p < 0.01$ compared to the control. B, Brain; L, Lung; S, Spleen; BF, Bursa of Fabricius.

The robust production of proinflammatory cytokines and chemokines such as *IL-1*β, *IL-6*, *IL-8*, *IL-10*, *TNF*-α, *MCP-1*, *IFN*-α, *IFN*-β, and *IFN*-γ in mammals during influenza virus infection, referred to as *cytokine storms*, have been confirmed to contribute to the severity of pathological damage via immune-mediated mechanisms (Walsh et al., 2011; Teijaro et al., 2014). In our study, the expression levels of *IL-1*β, *IL-10*, and *IFN*-β in the lungs and *MHC-II* in the brain were upregulated to a remarkably high level after infection with ZH283 and SW8, although were greater for ZH283 than SW8. Moreover, the expression level of S1PR1 in tested tissues following infection with ZH283 was less than that following infection with SW8. Consistent with the results of other studies (Walsh et al., 2011; Teijaro et al., 2014), our results demonstrated that the activation of S1PR1 can suppress the induction of cytokines, chemokines, and PRRs, meaning reducing morbidity and mortality, in chickens infected with H5N6. However, the specific mechanism of action remains to be determined.

In sum, both H5N6 AIVs were highly pathogenic to chickens, caused multiple systemic infections in tissues, and were efficiently and rapidly transmitted in chickens. Those results indicate that H5N6 viruses could be transmitted to domestic poultry, which represents a serious threat to the poultry industry and both human and animal health. Furthermore, the expression profiles of PRRs, proinflammatory cytokines, chemokines, and MHC molecules in the tested tissues of H5N6-infected chickens were involved in a tissue-dependent manner. Lastly, our experiments demonstrated that ZH283 was associated with greater pathogenicity in chickens, for high virus titers appeared in tested tissues early in the infection process and were accompanied by the excessive expression of cytokines. Such data provide new insights into the relationship between the pathogenicity of H5N6 AIVs and host immune responses to them in chickens.

AUTHOR CONTRIBUTIONS

SG, YK, and TR designed the study. YK, SG, RY, HM, ZW, BX, XD, FW, JX, and ML contributed reagents/materials and performed the statistical analysis. YK analyzed the data. YK and SG wrote the manuscript.

ACKNOWLEDGMENTS

This study was supported by in part by grant from the Science and Technology Projects of Guangdong Province (No. 2013B020224002), the poultry production technology of Guangdong system (No. 2016LM1115) and Shanxi Agriculture University research foundation for introducing doctor (No. 2013YJ14).

REFERENCES

Adams, S. C., Xing, Z., Li, J., and Cardona, C. J. (2009). Immune-related gene expression in response to H11N9 low pathogenic avian influenza virus infection in chicken and Pekin duck peripheral blood mononuclear cells. *Mol. Immunol.* 46, 1744–1749. doi: 10.1016/j.molimm.2009.01.025

Barton, G. M. (2007). Viral recognition by Toll-like receptors. *Semin. Immunol.* 19, 33–40. doi: 10.1016/j.smim.2007.01.003

Bi, Y., Chen, Q., Wang, Q., Chen, J., Jin, T., Wong, G., et al. (2016a). Genesis, evolution and prevalence of H5N6 avian influenza viruses in China. *Cell Host Microbe* 20, 810–821. doi: 10.1016/j.chom.2016.10.022

Bi, Y., Liu, H., Xiong, C., Di, L., Shi, W., Li, M., et al. (2016b). Novel avian influenza A (H5N6) viruses isolated in migratory waterfowl before the first human case reported in China, 2014. *Sci. Rep.* 6:29888. doi: 10.1038/srep29888

Bi, Y., Mei, K., Shi, W., Liu, D., Yu, X., Gao, Z., et al. (2015). Two novel reassortants of avian influenza A (H5N6) virus in China. *J. Gen. Virol.* 96(Pt 5), 975–981. doi: 10.1099/vir.0.000056

Butler, J., Stewart, C. R., Layton, D. S., Phommachanh, P., Harper, J., Payne, J., et al. (2016). Novel reassortant H5N6 influenza A virus from the Lao People's Democratic Republic is highly pathogenic in chickens. *PLoS ONE* 11:e0162375. doi: 10.1371/journal.pone.0162375

Claes, F., Morzaria, S. P., and Donis, R. O. (2016). Emergence and dissemination of clade 2.3.4.4 H5Nx influenza viruses-how is the Asian HPAI H5 lineage maintained. *Curr. Opin. Virol.* 16, 158–163. doi: 10.1016/j.coviro.2016.02.005

Cornelissen, J. B., Post, J., Peeters, B., Vervelde, L., and Rebel, J. M. (2012). Differential innate responses of chickens and ducks to low-pathogenic avian influenza. *Avian Pathol.* 41, 519–529. doi: 10.1080/03079457.2012.732691

Dhingra, M. S., Artois, J., Robinson, T. P., Linard, C., Chaiban, C., Xenarios, I., et al. (2016). Global mapping of highly pathogenic avian influenza H5N1 and H5Nx clade 2.3.4.4 viruses with spatial cross-validation. *Elife* 5:e19571. doi: 10.7554/eLife.19571

Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S., and Reis e Sousa, C. (2004). Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* 303, 1529–1531. doi: 10.1126/science.1093616

Du, Y., Chen, M., Yang, J., Jia, Y., Han, S., Holmes, E. C., et al. (2017). Molecular evolution and emergence of H5N6 avian influenza virus in central China. *J. Virol.* 91:e143-17. doi: 10.1128/JVI.00143-17

Gromme, M., and Neefjes, J. (2002). Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways. *Mol. Immunol.* 39, 181–202. doi: 10.1016/S0161-5890(02)00101-3

Hulse-Post, D. J., Franks, J., Boyd, K., Salomon, R., Hoffmann, E., Yen, H. L., et al. (2007). Molecular changes in the polymerase genes (PA and PB1) associated with high pathogenicity of H5N1 influenza virus in mallard ducks. *J. Virol.* 81, 8515–8524. doi: 10.1128/JVI.00435-07

Jhung, M. A., and Nelson, D. I. (2015). Outbreaks of avian influenza A (H5N2), (H5N8), and (H5N1) among birds–United States, December 2014-January 2015. *MMWR Morb. Mortal. Wkly. Rep.* 64:111.

Kang, Y., Liu, L., Feng, M., Yuan, R., Huang, C., Tan, Y., et al. (2017). Highly pathogenic H5N6 influenza A viruses recovered from wild birds in Guangdong, southern China, 2014–2015. *Sci. Rep.* 7:44410. doi: 10.1038/srep44410

Li, M., Zhao, N., Luo, J., Li, Y., Chen, L., Ma, J., et al. (2017). Genetic characterization of continually evolving highly pathogenic H5N6 influenza viruses in China, 2012-2016. *Front. Microbiol.* 8:260. doi: 10.3389/fmicb.2017.00260

Li, X., Fu, Y., Yang, J., Guo, J., He, J., Weng, S., et al. (2015). Genetic and biological characterization of two novel reassortant H5N6 swine influenza viruses in mice and chickens. *Infect. Genet. Evol.* 36, 462–466. doi: 10.1016/j.meegid.2015.08.017

Li, X., Shi, J., Guo, J., Deng, G., Zhang, Q., Wang, J., et al. (2014). Genetics, receptor binding property, and transmissibility in mammals of naturally isolated H9N2 Avian Influenza viruses. *PLoS Pathog.* 10:e1004508. doi: 10.1371/journal.ppat.1004508

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- ΔΔCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lowen, A. C., Mubareka, S., Steel, J., and Palese, P. (2007). Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* 3:1470–1476. doi: 10.1371/journal.ppat.0030151

MacDonald, M. R., Xia, J., Smith, A. L., and Magor, K. E. (2008). The duck toll like receptor 7: genomic organization, expression and function. *Mol. Immunol.* 45, 2055–2061. doi: 10.1016/j.molimm.2007.10.018

Medzhitov, R. (2001). Toll-like receptors and innate immunity. *Nat. Rev. Immunol.* 1, 135–145. doi: 10.1038/35100529

Pan, M., Gao, R., Lv, Q., Huang, S., Zhou, Z., Yang, L., et al. (2016). Human infection with a novel, highly pathogenic avian influenza A (H5N6) virus: virological and clinical findings. *J. Infect.* 72, 52–59. doi: 10.1016/j.jinf.2015.06.009

Pu, J., Wang, S., Yin, Y., Zhang, G., Carter, R. A., Wang, J., et al. (2015). Evolution of the H9N2 influenza genotype that facilitated the genesis of the novel H7N9 virus. *Proc. Natl. Acad. Sci. U.S.A.* 112, 548–553. doi: 10.1073/pnas.1422456112

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty percent endpoints. *Am. J. Epidemiol.* 27, 493–497. doi: 10.1093/oxfordjournals.aje.a118408

Rivera, J., Proia, R. L., and Olivera, A. (2008). The alliance of sphingosine-1-phosphate and its receptors in immunity. *Nat. Rev. Immunol.* 8, 753–763. doi: 10.1038/nri2400

Saito, T., Tanikawa, T., Uchida, Y., Takemae, N., Kanehira, K., and Tsunekuni, R. (2015). Intracontinental and intercontinental dissemination of Asian H5 highly pathogenic avian influenza virus (clade 2.3.4.4) in the winter of 2014-2015. *Rev. Med. Virol.* 25, 388–405. doi: 10.1002/rmv.1857

Shen, Y. Y., Ke, C. W., Li, Q., Yuan, R. Y., Xiang, D., Jia, W. X., et al. (2016). Novel Reassortant Avian Influenza A(H5N6) Viruses in Humans, Guangdong, China, 2015. *Emerg. Infect. Dis.* 22, 1507–1509. doi: 10.3201/eid2208.160146

Stephenson, I., Wood, J. M., Nicholson, K. G., Charlett, A., and Zambon, M. C. (2004). Detection of anti-H5 responses in human sera by HI using horse erythrocytes following MF59-adjuvanted influenza A/Duck/Singapore/97 vaccine. *Virus Res.* 103, 91–95. doi: 10.1016/j.virusres.2004.02.019

Sturm-Ramirez, K. M., Ellis, T., Bousfield, B., Bissett, L., Dyrting, K., Rehg, J. E., et al. (2004). Reemerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. *J. Virol.* 78, 4892–4901. doi: 10.1128/JVI.78.9.4892-4901.2004

Sun, H., Pu, J., Hu, J., Liu, L., Xu, G., Gao, G. F., et al. (2016). Characterization of clade 2.3.4.4 highly pathogenic H5 avian influenza viruses in ducks and chickens. *Vet. Microbiol.* 182, 116–122. doi: 10.1016/j.vetmic.2015.11.001

Teijaro, J. R., Walsh, K. B., Cahalan, S., Fremgen, D. M., Roberts, E., Scott, F., et al. (2011). Endothelial cells are central orchestrators of cytokine amplification during influenza virus infection. *Cell* 146, 980–991. doi: 10.1016/j.cell.2011.08.015

Teijaro, J. R., Walsh, K. B., Rice, S., Rosen, H., and Oldstone, M. B. (2014). Mapping the innate signaling cascade essential for cytokine storm during influenza virus infection. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3799–3804. doi: 10.1073/pnas.1400593111

Tscherne, D. M., and Garcia-Sastre, A. (2011). Virulence determinants of pandemic influenza viruses. *J. Clin. Invest.* 121, 6–13. doi: 10.1172/JCI44947

Vanderven, H. A., Petkau, K., Ryan-Jean, K. E., Aldridge, J. R. Jr., Webster, R. G., and Magor, K. E. (2012). Avian influenza rapidly induces antiviral genes in duck lung and intestine. *Mol. Immunol.* 51, 316–324. doi: 10.1016/j.molimm.2012.03.034

Walsh, K. B., Teijaro, J. R., Wilker, P. R., Jatzek, A., Fremgen, D. M., Das, S. C., et al. (2011). Suppression of cytokine storm with a sphingosine analog provides protection against pathogenic influenza virus. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12018–12023. doi: 10.1073/pnas.1107024108

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56, 152–179.

Wei, L., Jiao, P., Song, Y., Cao, L., Yuan, R., Gong, L., et al. (2013). Host immune responses of ducks infected with H5N1 highly pathogenic avian influenza viruses of different pathogenicities. *Vet. Microbiol.* 166, 386–393. doi: 10.1016/j.vetmic.2013.06.019

WHO (2016). *Cumulative Number of Confirmed Human Cases for Avian Influenza A(H5N1) Reported to WHO, 2003-2016*. Available at: http://www.who.int/influenza/human_animal_interface/2016_10_03_tableH5N1.pdf?ua=1

Williams, A., Peh, C. A., and Elliott, T. (2002). The cell biology of MHC class I antigen presentation. *Tissue Antigens* 59, 3–17. doi: 10.1034/j.1399-0039.2002.590103.x

Yu, Z., Gao, X., Wang, T., Li, Y., Xu, Y., Chu, D., et al. (2015). Fatal H5N6 avian influenza virus infection in a domestic cat and wild birds in China. *Sci. Rep.* 5:10704. doi: 10.1038/srep10704

Yuan, R., Cui, J., Zhang, S., Cao, L., Liu, X., Kang, Y., et al. (2014). Pathogenicity and transmission of H5N1 avian influenza viruses in different birds. *Vet. Microbiol.* 168, 50–59. doi: 10.1016/j.vetmic.2013.10.013

Yuan, R., Wang, Z., Kang, Y., Wu, J., Zou, L., Liang, L., et al. (2016). Continuing reassortant of H5N6 subtype highly pathogenic avian influenza virus in Guangdong. *Front. Microbiol.* 7:520. doi: 10.3389/fmicb.2016.00520

Zhang, P., Tang, Y., Liu, X., Liu, W., Zhang, X., Liu, H., et al. (2009). A novel genotype H9N2 influenza virus possessing human H5N1 internal genomes has been circulating in poultry in eastern China since 1998. *J. Virol.* 83, 8428–8438. doi: 10.1128/JVI.00659-09

Zhang, P., Tang, Y., Liu, X., Peng, D., Liu, W., Liu, H., et al. (2008). Characterization of H9N2 influenza viruses isolated from vaccinated flocks in an integrated broiler chicken operation in eastern China during a 5 year period (1998-2002). *J. Gen. Virol.* 89(Pt 12), 3102–3112. doi: 10.1099/vir.0.2008/005652-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gao, Kang, Yuan, Ma, Xiang, Wang, Dai, Wang, Xiao, Liao and Ren. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 July 2016
doi: 10.3389/fmicb.2016.01066

Pathogenesis and Phylogenetic Analyses of Two Avian Influenza H7N1 Viruses Isolated from Wild Birds

Hongmei Jin[1†‡], Deli Wang[2†‡], Jing Sun[1,3†‡], Yanfang Cui[2], Guang Chen[1,4], Xiaolin Zhang[1], Jiajie Zhang[1], Xiang Li[1], Hongliang Chai[1], Yuwei Gao[5]*, Yanbing Li[2]* and Yuping Hua[1]**

[1] College of Wildlife Resources, Northeast Forestry University, Harbin, China, [2] Harbin Veterinary Research Institute, Chinese Academy of Agriculture Sciences, Harbin, China, [3] Research Institute of Forestry Ecology, Environment and Protection, Beijing, China, [4] Hubei Province Wildlife Epidemic Disease Center, Wuhan, China, [5] Changchun Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Changchun, China

The emergence of human infections with a novel H7N9 influenza strain has raised global concerns about a potential human pandemic. To further understand the character of other influenza viruses of the H7 subtype, we selected two H7N1 avian influenza viruses (AIVs) isolated from wild birds during routine surveillance in China: A/Baer's Pochard/Hunan/414/2010 (BP/HuN/414/10) (H7N1) and A/Common Pochard/Xianghai/420/2010 (CP/XH/420/10) (H7N1). To better understand the molecular characteristics of these two isolated H7N1 viruses, we sequenced and phylogenetically analyzed their entire genomes. The results showed that the two H7N1 strains belonged to a Eurasian branch, originating from a common ancestor. Phylogenetic analysis of their hemagglutinin (HA) genes showed that BP/HuN/414/10 and CP/XH/420/10 have a more distant genetic relationship with A/Shanghai/13/2013 (H7N9), with similarities of 91.6 and 91.4%, respectively. To assess the replication and pathogenicity of these viruses in different hosts, they were inoculated in chickens, ducks and mice. Although, both CP/XH/420/10 and BP/HuN/414/10 can infect chickens, ducks and mice, they exhibited different replication capacities in these animals. The results of this study demonstrated that two low pathogenic avian influenza (LPAI) H7N1 viruses of the Eurasian branch could infect mammals and may even have the potential to infect humans. Therefore, it is important to monitor H7 viruses in both domestic and wild birds.

Keywords: H7N1, H7N9, Avian influenza virus, phylogenetic analysis, pathogenic analyses

OPEN ACCESS

Edited by:
Aeron Hurt,
WHO Collaborating Centre for Reference and Research on Influenza, Australia

Reviewed by:
Jeff Michael Butler,
Commonwealth Scientific and Industrial Research Organisation-Australian Animal Health Laboratory, Australia
Shailesh D. Pawar,
National Institute of Virology, India
Karoline Bragstad,
Norwegian Institute of Public Health, Norway

***Correspondence:**
Hongliang Chai
17758625@163.com;
Yuwei Gao
gaoyuwei@gmail.com;
Yanbing Li
lyb@hvri.ac.cn;
Yuping Hua
yuping_hua@126.com

[†] These authors have contributed equally to this work.
[‡] Co-senior authors.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *11 March 2016*
Accepted: *24 June 2016*
Published: *07 July 2016*

Citation:
Jin H, Wang D, Sun J, Cui Y, Chen G, Zhang X, Zhang J, Li X, Chai H, Gao Y, Li Y and Hua Y (2016) Pathogenesis and Phylogenetic Analyses of Two Avian Influenza H7N1 Viruses Isolated from Wild Birds. Front. Microbiol. 7:1066. doi: 10.3389/fmicb.2016.01066

INTRODUCTION

Currently, avian influenza outbreaks and epidemics, particularly those of the H5 or H7 subtype, result in huge economic losses to the poultry industry and pose a serious threat to human health (Senne et al., 1996). Over the past two decades, many infections with influenza virus subtype H7 have occurred. For example, between 1999 and 2000, a H7N1 virus outbreak in Italy resulted in the death of more than 13 million chickens and caused extensive economic losses (Capua et al., 2003). In 2003, 30 million birds were culled in the Netherlands, Belgium, and Germany after an H7N7 subtype influenza outbreak. During that outbreak, the H7N7 subtype avian influenza virus

(AIV) also infected 89 people, causing conjunctivitis with one fatal case (Fouchier et al., 2004). In February 2013, human infections with a novel H7N9 AIV were first reported in China and caused a widespread public health concern. By the end of March 2016, the H7N9 virus had infected 619 people and caused 255 fatalities in China (http://www.nhfpc.gov.cn/zwgkzt/ptggg/list.shtml). In addition, the emergence of the novel human H7N9 LPAI viruses in 2013 (Yu et al., 2015) demonstrated that H7 LPAI viruses can infect humans. However, wild aquatic birds, particularly waterfowl, waders and gulls, are regarded as a major natural reservoir of LPAI viruses, and these birds generally remain healthy while carrying the viruses (Alexander, 2000). Therefore, strengthening AIV surveillance in water birds is required.

In this study, we isolated a strain of H7N1 AIV from a healthy Baer's Pochard during AIV surveillance in Hunan Province in 2010. The same year, we also isolated a strain of H7N1 AIV from a healthy Common Pochard in Jilin Province. To better understand the molecular and biological properties of the two strains, we performed phylogenetic analysis based on complete genomic sequence data and assessed the replication and pathogenic potential of the two isolated viruses in ducks, chickens and mice. We also analyzed the receptor binding characteristics of the two H7 isolates. These studies expand our understanding of the latent evolutionary and transmission features of the H7 subtype viruses and would aid in disease control and pandemic preparedness efforts.

MATERIALS AND METHODS

Ethics Statements

All animal studies were approved by the Institutional Animal Care and Use Committee of the Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences. All animal procedures were carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the Ministry of Science and Technology of the People's Republic of China. All experiments were performed in a biosafety level 2+ laboratory (enhanced animal biosafety level 2 laboratory and a negative pressure-ventilation laboratory) at Harbin Veterinary Research Institute (Harbin, China).

Viruses Used in the Study

In our study, the avian influenza virus strain, A/chicken/Jilin/HU/02 (H5N1), which can specifically bind to α-2,3-linked sialic acid receptors, and the human influenza virus strain A/Jilin/31/2005 (H1N1), which can specifically bind to α-2,6-linked sialic acid receptors, were used. The strains were stored at Changchun Veterinary Research Institute, Chinese Academy of Agricultural Sciences. A/Baer's Pochard/Hunan/414/2010 (BP/HuN/414/10) (H7N1) and A/Common Pochard/Xianghai/420/2010 (CP/XH/420/10) (H7N1) isolated from a Baer's Pochard in Hunan Province and a Common Pochard in Jilin Province, respectively, were used in this study.

Virus Isolation

The two viruses analyzed in this study, BP/HuN/414/10 (H7N1) and CP/XH/420/10 (H7N1) were isolated during avian influenza surveillance in 2010. The method used to isolate these viruses involved oropharyngeal and cloacal swabs or fecal samples being suspended in antibiotic (1000 ug/ml penicillin and streptomycin)-treated phosphate-buffered saline (PBS) and centrifuged at 5000 rpm for 10 min at 4°C. The allantoic cavities of 10-day-old embryonated specific-pathogen free (SPF) chicken eggs were inoculated with the supernatant of oropharyngeal and cloacal swabs or fecal samples. The presence of AIV was

TABLE 1 | Selected characteristic amino acids of H7N1 subtype AIVs isolated from wild birds.

Gene	Key residues	Comments	The isolates		
			CP/XH/420/10 (H7N1)	BP/HuN/414/10 (H7N1)	A/Shanghai/13/2013 (H7N9)
HA (H3 numbering)	Cleavage sites	Characteristic of low-pathogenic in chickens (Steinhauer, 1999)	PELPKGR↓GLFGAI	PELPKGR↓GLFGAI	PEIPKGR↓GLFGAI
	Ser138Ala(S-A)	Favor mammalian adaptation (Fan et al., 2015)	A	A	A
	Thr160Ala(T-A)	Increased affinity toward α-2,6-linked sialic acid receptor (Fan et al., 2015)	A	A	A
	Gln226Leu(Q-L)	Increased affinity toward α-2,6-linked sialic acid receptor (Yamada et al., 2006)	Q	Q	I
	Gly228Ser(G-S)	Increased affinity toward α-2,6-linked sialic acid receptor (Yamada et al., 2006)	G	G	G
PB2	Glu627Lys(E-K)	Mammalian adaptation (Li et al., 2005)	E	E	K
	Asp701Asn(D-N)	Mammalian adaptation (Li et al., 2005)	D	D	D
M2	Ser31Asn(S-N)	Amantadine resistance (Lee et al., 2008)	S	S	N
NS1	Pro42Ser(P-S)	Increased virulence in mice (Jiao et al., 2008)	S	S	S

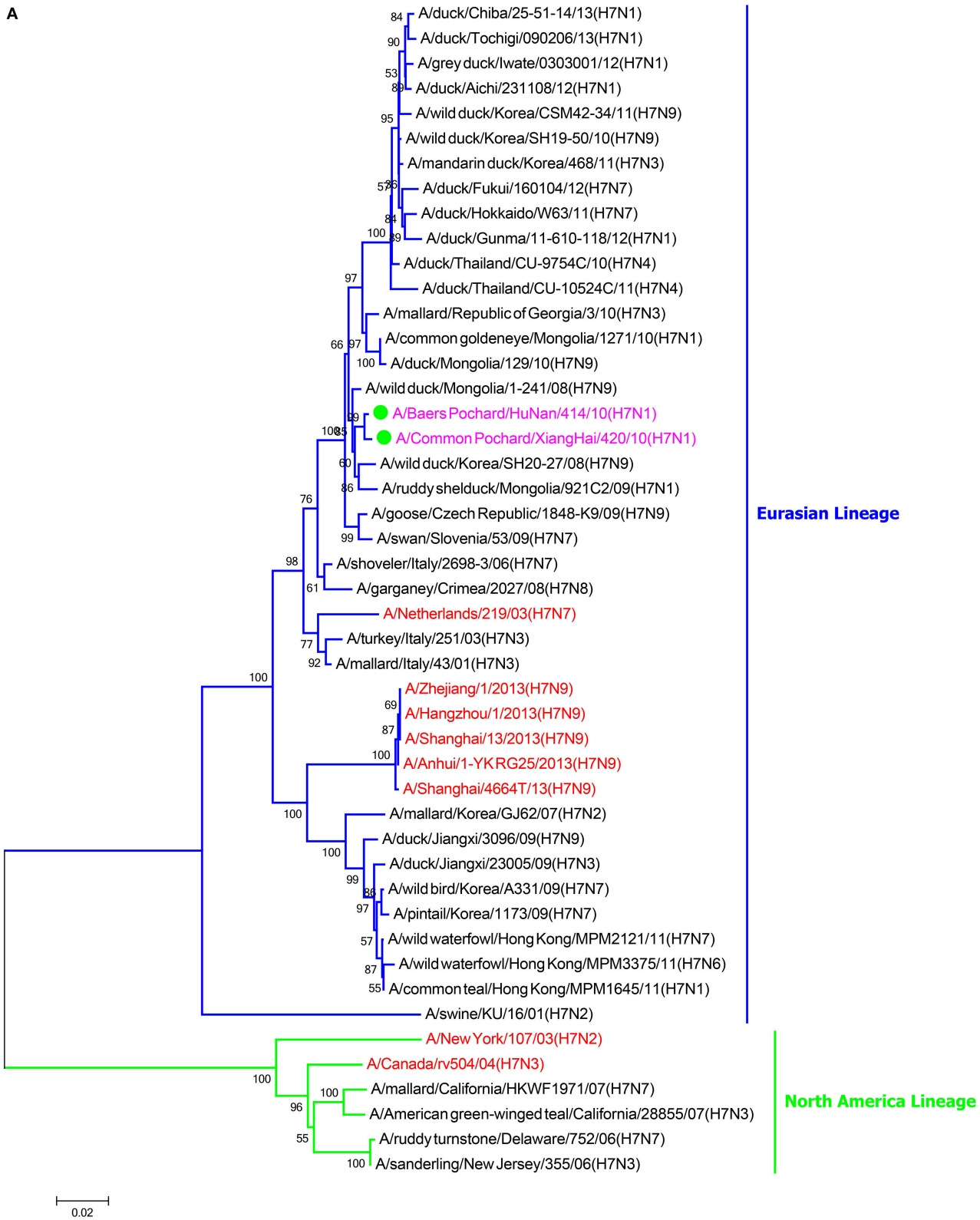


FIGURE 1 | Continued

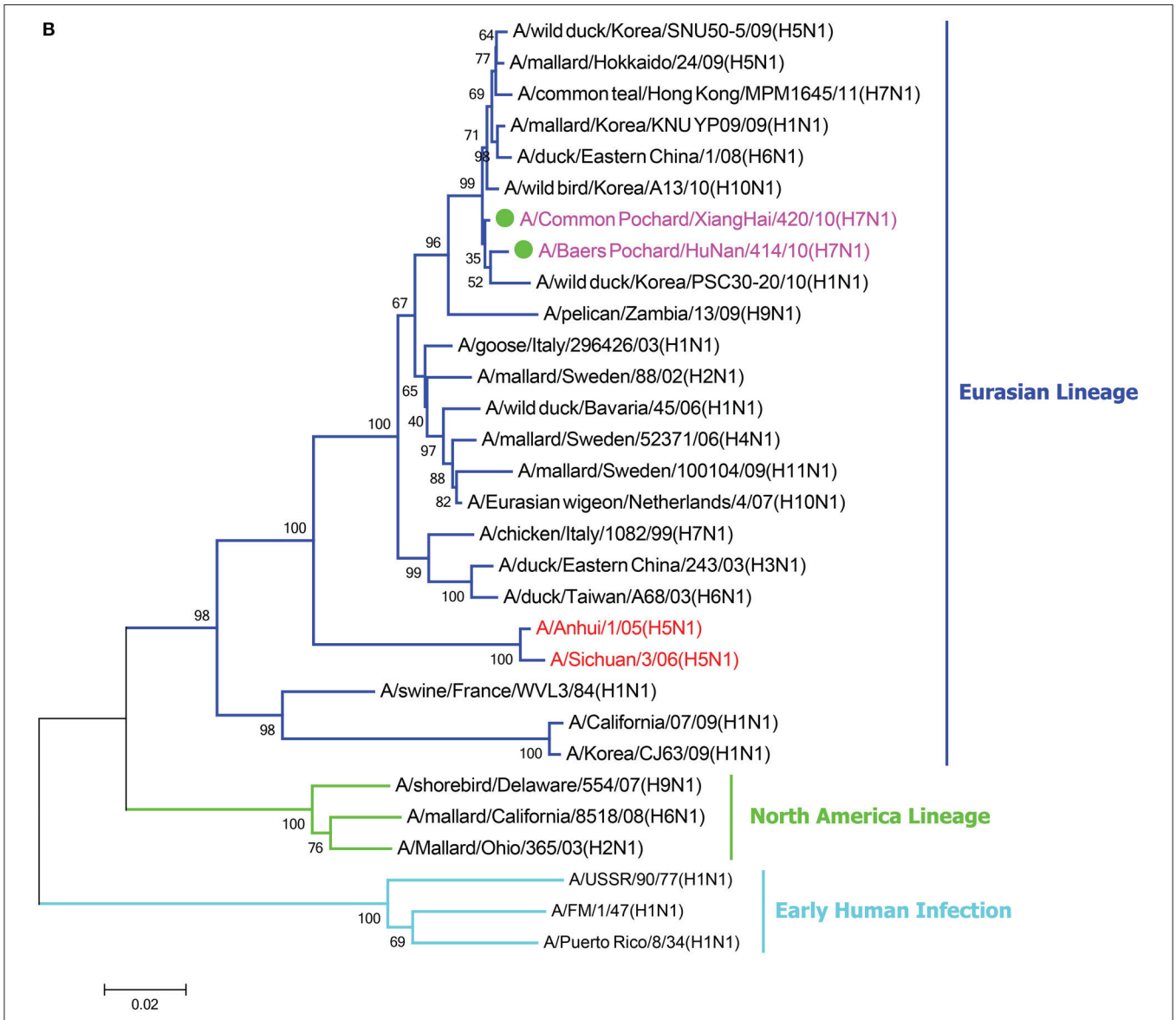


FIGURE 1 | Continued

confirmed with RT-PCR (See Supplementary Table 2). The 50% egg infection dose (EID_{50}) was calculated according to the method described by Reed and Muench (1938). The allantoic fluid of the purified virus was stored at −80°C until use. All procedures were performed under aseptic conditions.

Genome Sequencing and Phylogenetic Analysis

Viral RNA was extracted from the allantoic fluid using TRIZOL Reagent (Invitrogen Carlsbad, CA, USA) and reverse transcribed using the primer 5′-AGCRAAAGCAGG-3′. The PCR products of eight fragments of the H7N1 virus were sequenced with a set of specific sequencing primers. The sequence data were compiled with the SEQMAN program (DNASTAR, Madison, WI). All reference sequences used in this study were obtained from the National Center for Biotechnology Information (NCBI) GenBank database and the Global Initiative on Sharing All Influenza Data (See Supplementary Table 1). DNASTAR's MegAlign was used to perform the sequence homology and key amino acid site analyses. MEGA 6.06 software was used to perform multiple sequence alignments with the Clustal W algorithm, and a phylogenetic tree was generated with the neighbor-joining method and bootstrap test (1000 replicates) based on the sequences for the open reading frames. Potential glycosylation sites were analyzed with the NetNGlyc 1.0 online software (www.cbs.dtu.dk/services/NetNGlyc/). GenBank accession numbers are JQ973643-JQ973650 for BP/HuN/414/10 and KU663402-KU663409 for CP/XH/420/10.

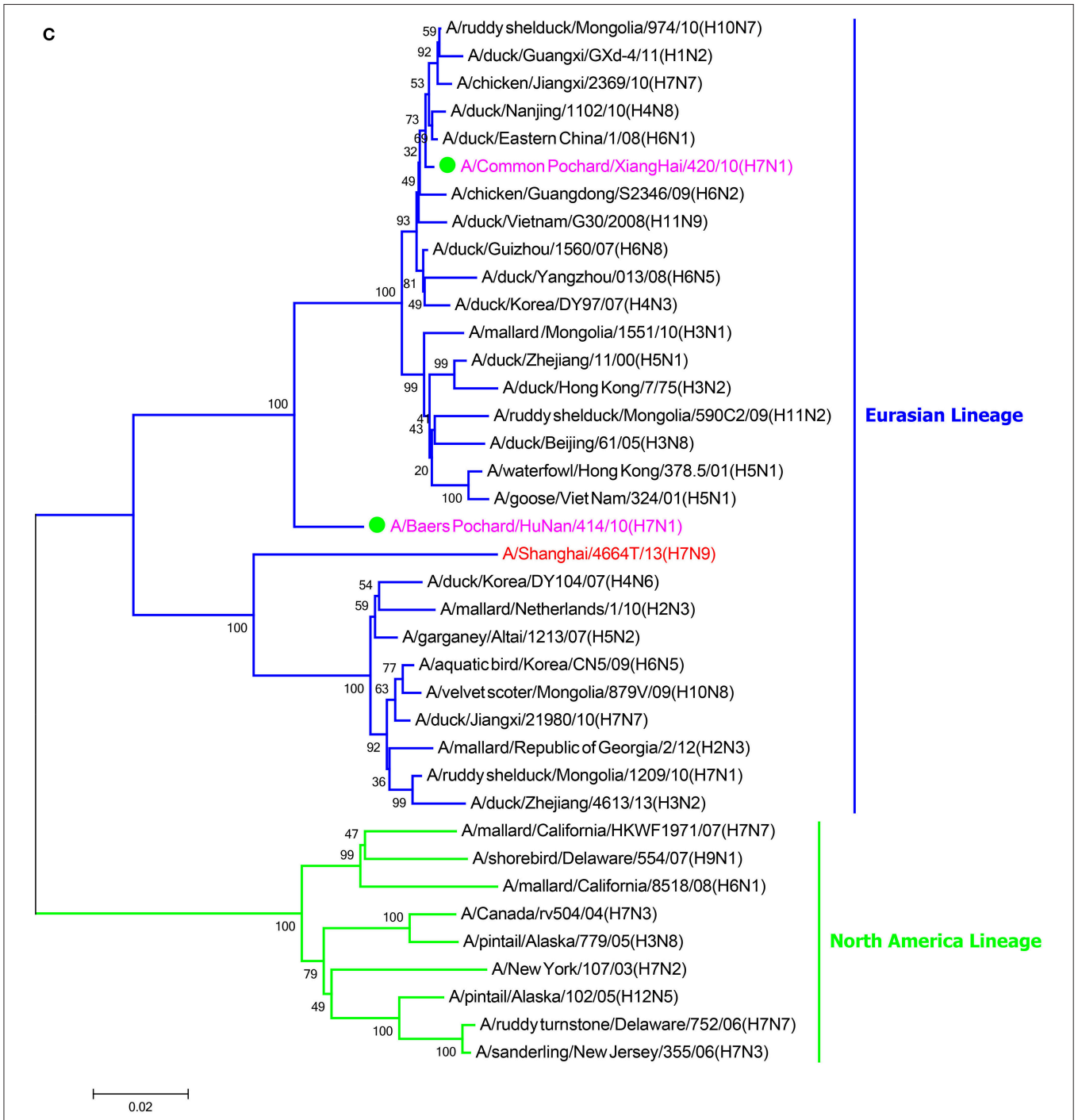


FIGURE 1 | Continued

Infection of Chickens and Ducks

To examine the replication and transmission of the two isolated H7N1 viruses in chickens and ducks, two groups of 4-week-old SPF chickens (white leghorn) and 3-week-old SPF ducks (shelduck) (eight birds/group) were inoculated with 100 μl 10^6 EID_{50} virus in each bird. Twenty-four hours after inoculation, three additional chickens and ducks were placed in the same isolation units to monitor contact infection (Fan et al., 2015). All birds were monitored every day for clinical symptoms or death until 21 days post infection (dpi). Viral shedding was monitored through sampling of oropharyngeal and cloacal from infected and contacted birds at 3, 5, and 7 dpi. At 3 dpi, three birds in each group were euthanized, and tissues, including the brain, kidney, spleen, lung, bursa, trachea, cecal tonsil, thymus, and

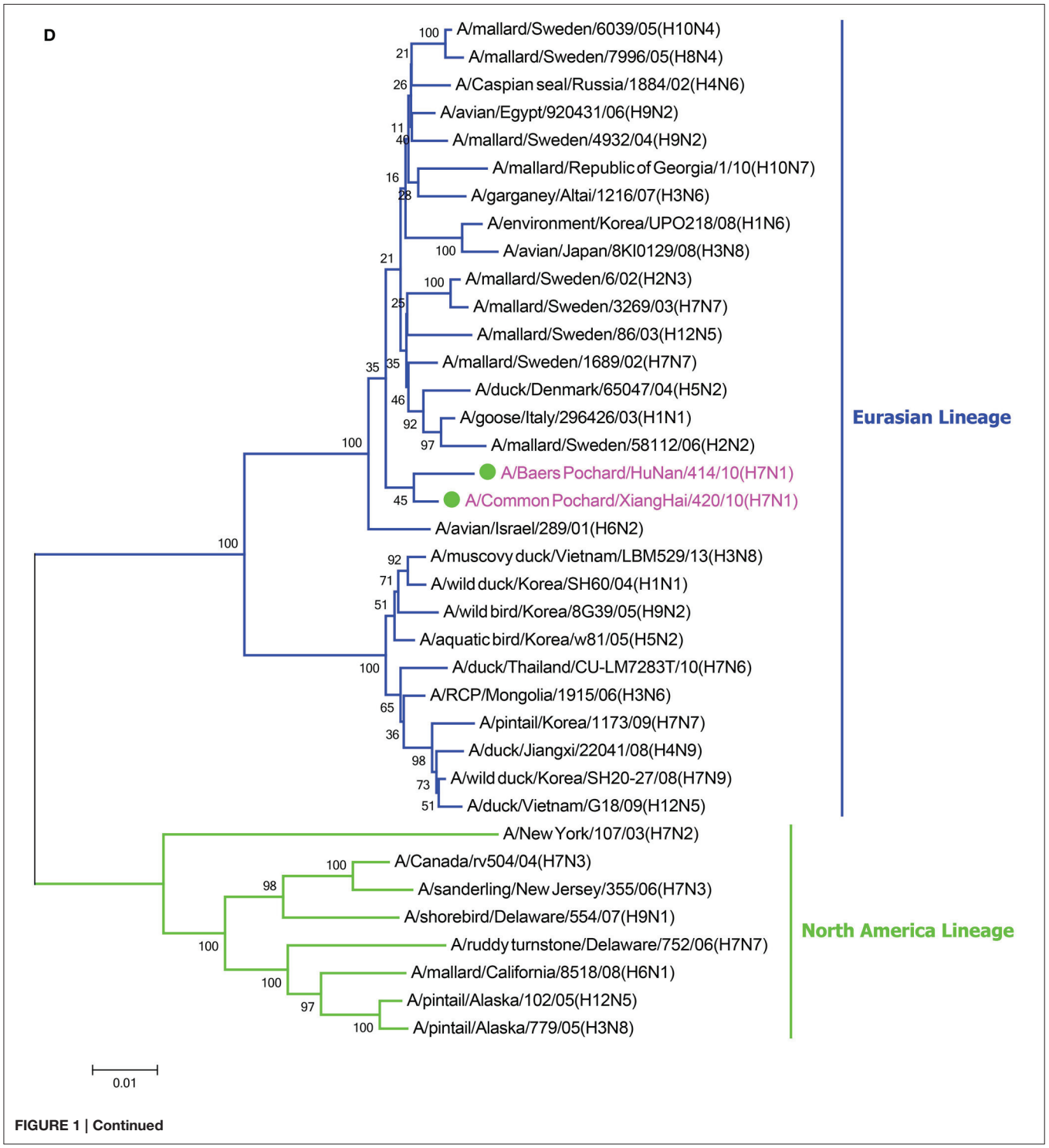


FIGURE 1 | Continued

pancreas, were collected aseptically for virus titration. Serum samples were collected from each bird at 14 and 21 dpi for detection of antibodies using the HI assay.

Infection of Mice

To investigate the virulence of the two H7N1 viruses in mice, two groups (86-week-old female BALB/c mice/group) were lightly anesthetized with CO_2 and inoculated intranasally with 50 μl of 10^6 EID_{50} of the H7N1 influenza virus (Chen et al., 2004). At 3 dpi, three inoculated mice were euthanized with a peritoneal injection of sodium pentobarbital at a dose of 200 mg/kg, and their organs, including the lung, kidney, spleen, turbinate, and brain, were collected for viral titration and histopathological evaluation. The titers for virus infectivity in

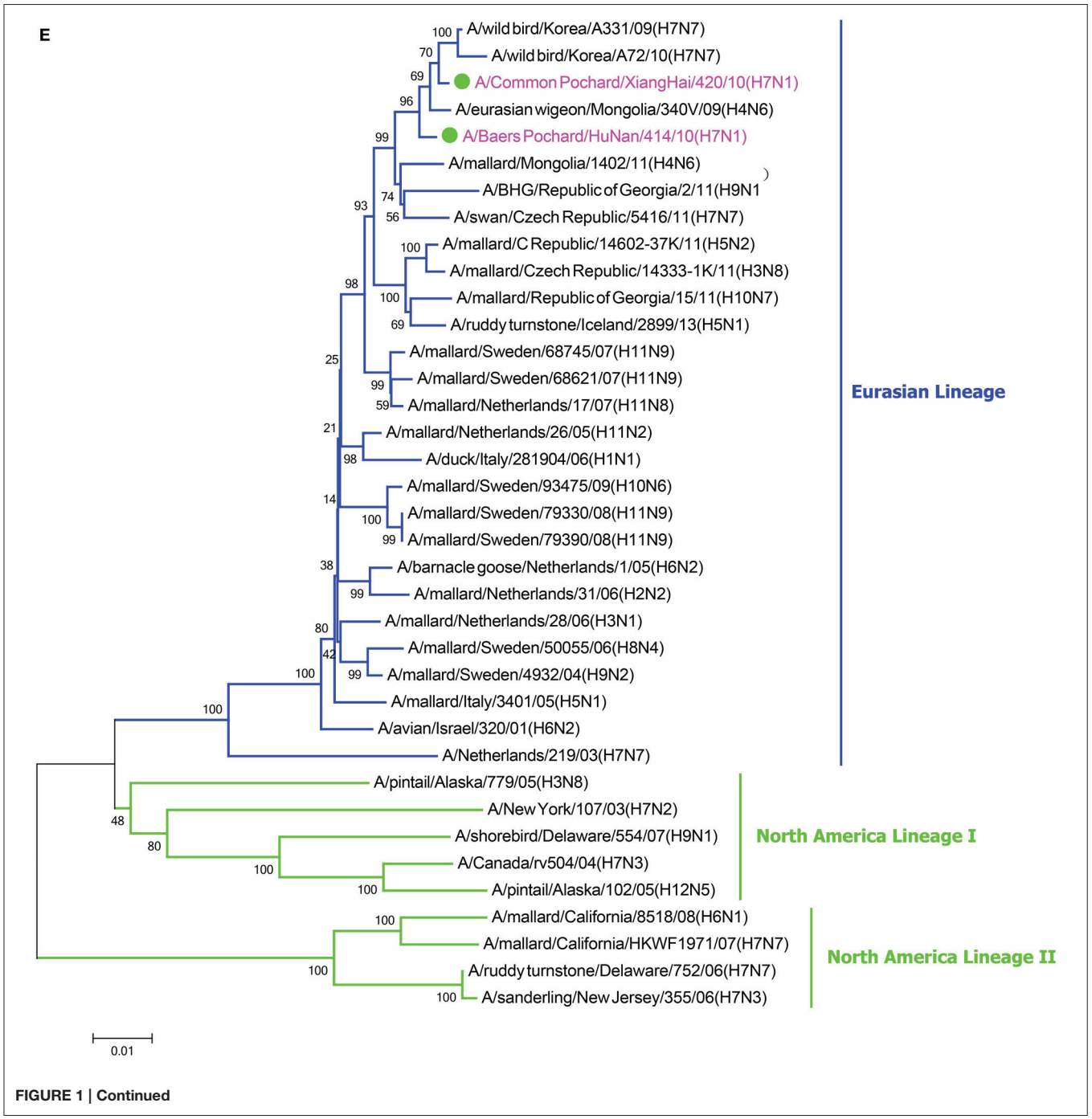


FIGURE 1 | Continued

eggs were calculated with the method described by Reed and Muench (1938) and Li et al. (2005). The remaining five mice in each group were monitored daily for 14 days for weight loss and mortality (Zhao et al., 2012). The control group (five mice) was mock infected with PBS and monitored daily for 14 days for weight loss and mortality. All weight changes were calculated based on the average of five mice using the percent of the everyday weight divided by the first-day weight and multiplied by 100 (Ye et al., 2016). Significant changes in body weight were calculated by one-way ANOVA, and $P < 0.05$ was considered statistically significant (Mancinelli et al., 2016).

Analysis of Receptor Specificity of the Two Strains of H7N1 Virus

To prepare the red blood cell suspension, Alsever's solution anticoagulant was added at a dilution of 1:1 upon collection of chicken and sheep red blood cells. The chicken and sheep red blood cells were washed three times with PBS, centrifuged at 2000 rpm for 5 min at 4°C each time, and adjusted to final

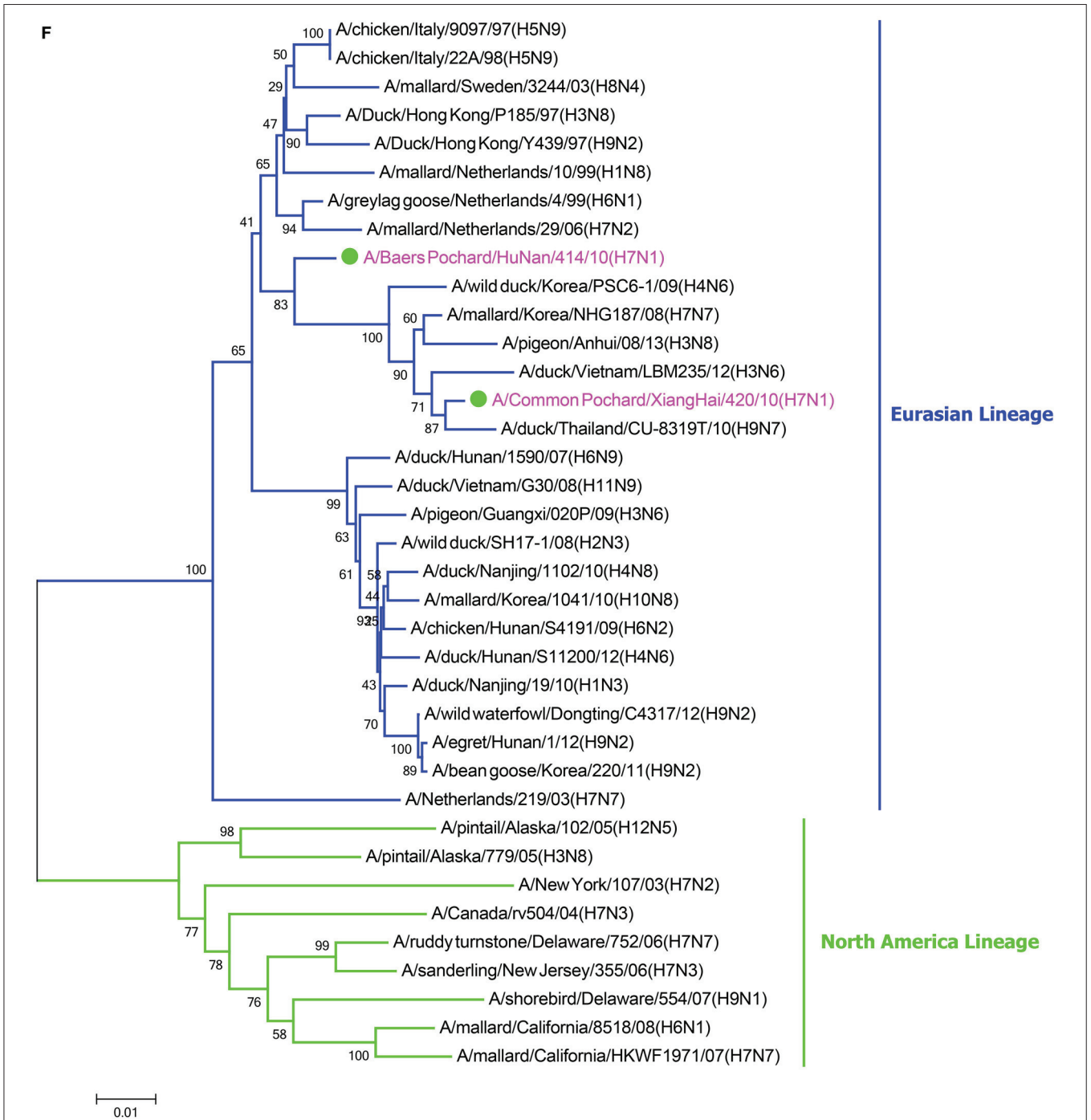


FIGURE 1 | Continued

working concentrations (10 and 1%, respectively) with PBS and stored at 4°C.

For the sialidase treatment, 90 μl of a 10% suspension of chicken red blood cells was treated with 10 μl of α-2,3-sialidase (50 mU/μl) (TaKaRa, Dalian, China) for 10 min at 37°C. The sample was then washed two times with PBS, centrifuged at 2000 rpm for 5 min at 4°C each time, adjusted to a final working concentration (0.75%) with PBS, and stored at 4°C. The chicken red blood cells were treated with α-2,3-sialidase to eliminate all receptors except for the α-2,6-linked sialic acid receptor.

For the vibrio cholera neuraminidase (VCNA) (TaKaRa, Dalian, China) treatment, 90 μl of a 10% suspension of chicken red blood cells was treated with 10 μl of VCNA (50 mU/μl) for 1 h at 37°C, washed three times with PBS, centrifuged at 2000

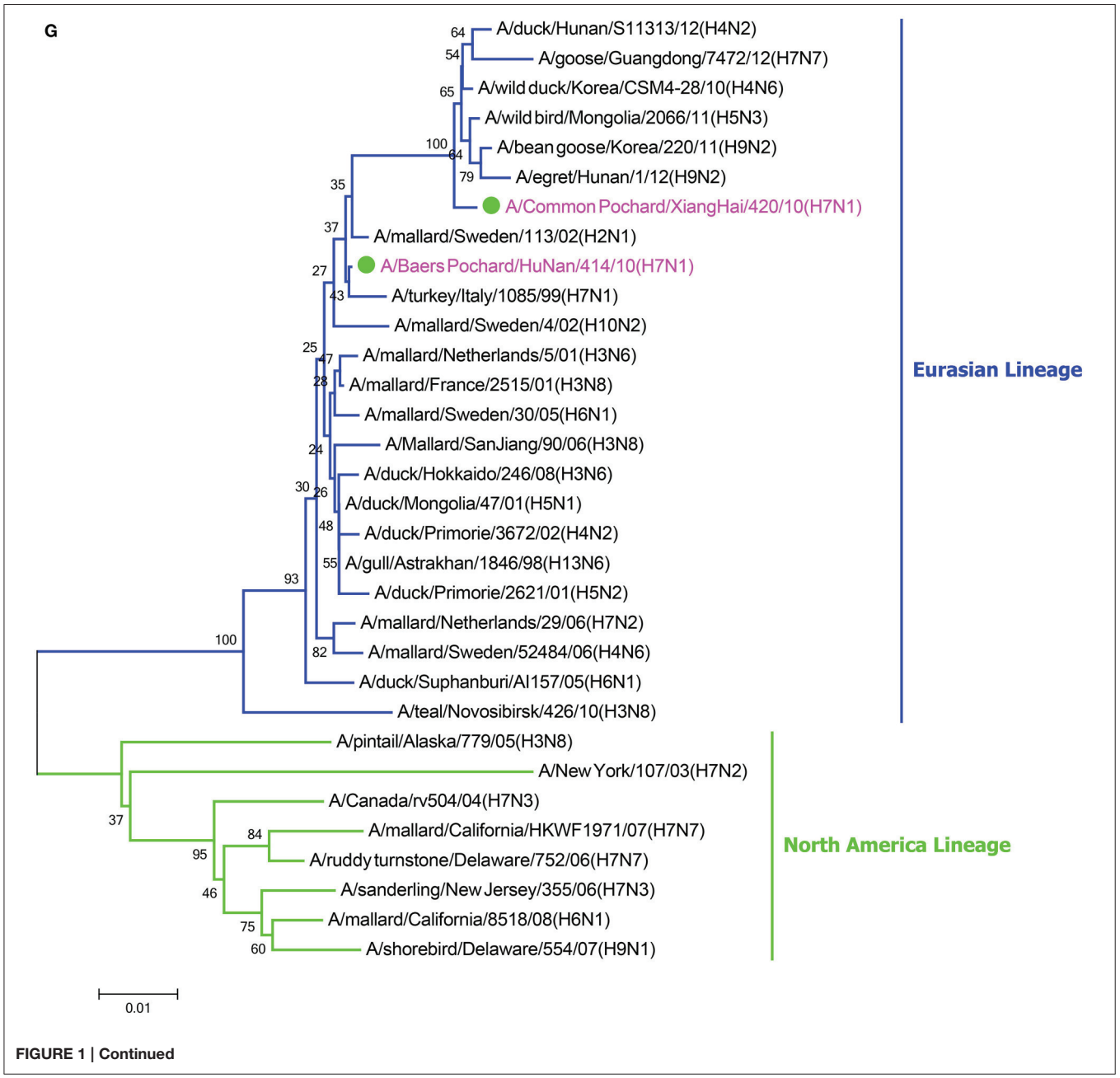


FIGURE 1 | Continued

rpm for 5 min at 4°C each time, adjusted to a final working concentration (0.75%) with PBS, and stored at 4°C. The chicken red blood cells were treated with VCNA to eliminate the α-2,3-linked sialic acid receptors and α-2,6-linked sialic acid receptors.

Both the 10% chicken red blood cells (with α-2,3-linked sialic acid receptors and α-2,6-linked sialic acid receptors) and 1% sheep red blood cells (with only α-2,3-linked sialic acid receptors) were then diluted to a concentration of 0.75%. Four virus suspensions, including BP/HuN/414/10, CP/XH/420/10, A/Jilin/31/2005 (H1N1), and A/chicken/Jilin/HU/02 (H5N1) were subsequently diluted with PBS to a dilution of 1:32, and the agglutination of red blood cells caused by diluted BP/HuN/414/10, CP/XH/420/10, A/Jilin/31/2005 (H1N1), and A/chicken/Jilin/HU/02 (H5N1) was determined using sheep red blood cells (0.75%), chicken red blood cells (0.75%), chicken red blood cells treated with a-2,3-sialidase (0.75%), and chicken red blood cells treated with VCNA (0.75%), respectively (Sun et al., 2008). This experiment was completed at the Changchun Veterinary Research Institute, Chinese Academy of Agricultural Sciences.

RESULTS

Mutation Analysis

Q226L and G228S mutations were not detected in the hemagglutinin (HA) protein, which indicates that the two H7N1

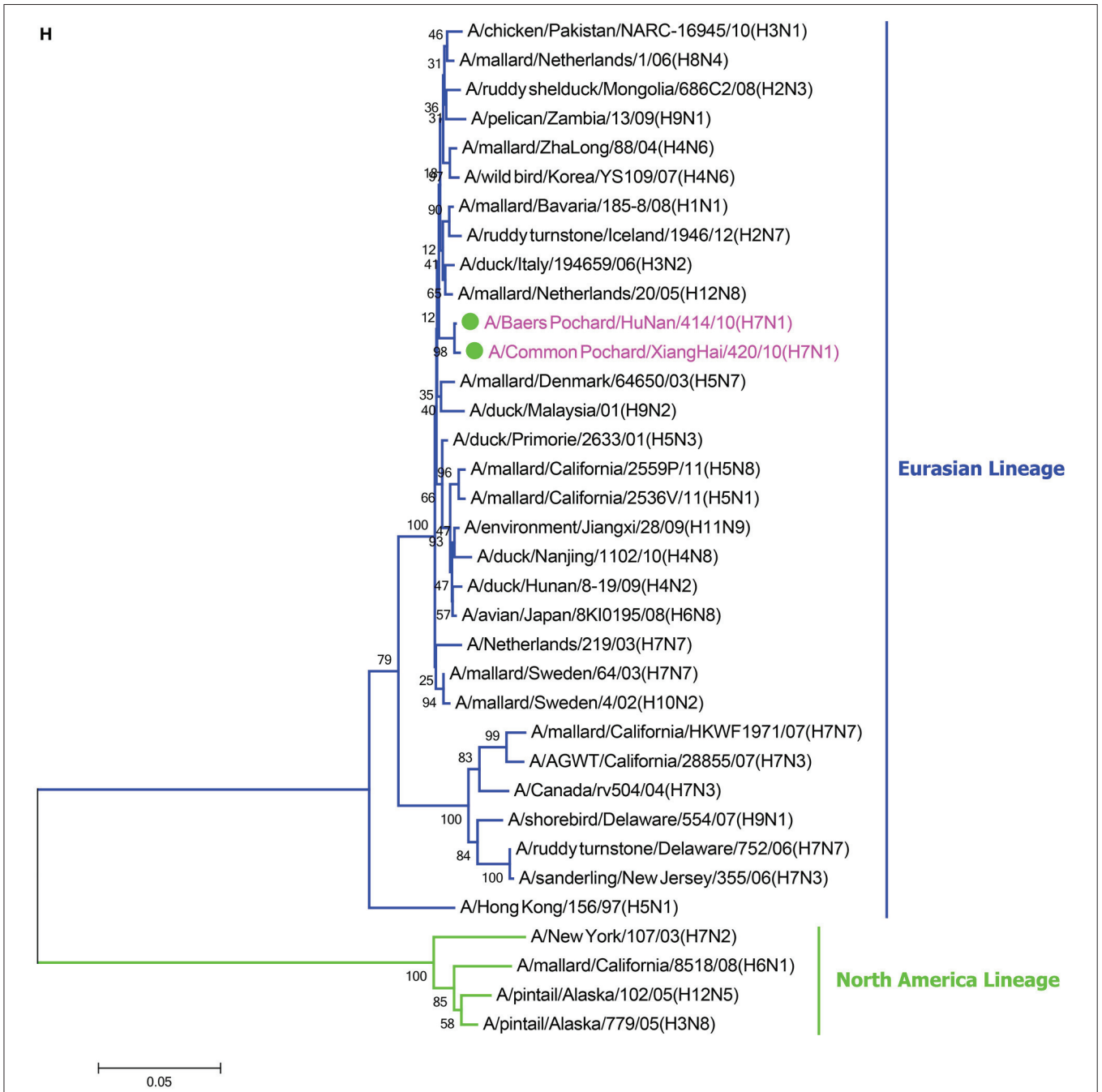


FIGURE 1 | A phylogenetic tree based on the open reading frame sequences of the HA (A), NA (B), PB2 (C), PB1 (D), PA (E), NP (F), M (G), and NS (H) gene segments of two H7N1 AIVs. The two isolated H7N1 viruses are indicated by a closed circle and highlighted in purple. Human isolates are highlighted in red. Different lineages are highlighted in different colors. The trees were constructed using the neighbor-joining method of MEGA6.0 6 with 1000 bootstrap trials to assign confidence to the groupings.

viruses may retain the characteristic of preferential binding to avian-like α-2,3-linked sialic acid receptors (Stevens et al., 2006; Yamada et al., 2006). Q226L and G228S mutations in the HA protein were not detected in the A/Shanghai/13/2013 (H7N9) strain, whereas HA S138A and T160A mutations were found in the two H7N1 viruses as well as the A/Shanghai/13/2013 (H7N9) strain. The two isolated viruses showed no E627K and D701N mutations in the PB2 protein, which plays an important role in the adaptation of AIVs to mammals (Katz et al., 2000; Li et al., 2005), but the E627K mutation was detected in the A/Shanghai/13/2013 (H7N9) strain. The amino acid substitution S31N was not detected in the M2 protein, indicating that these viral strains are sensitive to amantadine inhibitors (Lee et al., 2008), but it was detected in the A/Shanghai/13/2013

(H7N9) strain. The two H7N1 viruses and A/Shanghai/13/2013 (H7N9) exhibited mutations at position P42S of the NS1 protein (**Table 1**), which can increase virulence in mice (Jiao et al., 2008).

Phylogenetic Analysis

To clarify the genetic relationship of the two H7N1 viruses, we sequenced the entire genome of each virus and compared the eight gene segments of each virus with sequences of typical influenza viruses obtained from the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/genbank/) and Global Initiative on Sharing All Influenza Data (http://platform.gisaid.org).

In the HA phylogenetic tree (**Figure 1A**), the two viruses clustered into the Eurasian branch, and both viruses shared a close genetic relationship with 99.3% nucleotide identity in the HA gene, indicating that the HA genes of the two viruses likely originated from the same source. Both BP/HuN/414/10 and CP/XH/420/10 were most closely related to the A/wild duck/Mongolia/1-241/2008 (H7N9) strain, with 97.9 and 98.3% nucleotide identities, respectively (**Table 2**). However, a more distant genetic relationship with the A/Shanghai/13/2013 (H7N9) strain was observed, with 91.6 and 91.4% nucleotide identities, respectively. As shown in the NA phylogenetic tree (**Figure 1B**), both H7N1 isolates belong to the Eurasian lineage and share the greatest sequence homology (99.2 and 99.4%) with the A/wild bird/Korea/A13/2010 (H10N1) strain (**Table 2**). The highest nucleotide sequence identities relative to the internal gene fragments of the two isolated H7N1 AIVs and the phylogenetic trees of the two isolated H7N1 influenza viruses' internal genes are shown in **Table 2**, **Figures 1C–H**. Remarkably, the data showed that the constellation of the internal genome segments of the two H7N1 viruses were substantially different. While the PB1, PB2, PA, NP, M, and NS genes of CP/XH/420/10 originated from H9N2, H4N8, H7N7, H7N7, H4N6, and H4N6 viruses, respectively, the PB1, PB2, PA, NP, M, and NS genes of BP/HuN/414/10 originated from H9N2, H4N8, H7N7, H5N9, H2N1, and H1N1 viruses, respectively. These results indicated that the two H7N1 strains have arisen through reassortment with different progenitor viruses.

Chicken and Duck Experiments

Throughout the experiment, no chickens or ducks showed any clinical symptoms or mortality. We did not detect any

TABLE 2 | The highest nucleotide identity of the whole genomes of two H7N1 influenza viruses.

Gene segment	BP/HuN/414/10(H7N1)			CP/XH/420/10(H7N1)		
	Closest viruses	Nucleotide identity (%)	Accession No.	Closest viruses	Nucleotide identity (%)	Accession No.
HA	A/wild duck/Mongolia/1-241/2008(H7N9)	99	JN029686.1	A/wild duck/Mongolia/1-241/2008(H7N9)	98.9	JN029686.1
NA	A/wild bird/Korea/A13/2010(H10N1)	99.2	JN817546.1	A/wild bird/Korea/A13/2010(H10N1)	99.4	JN817546.1
PB2	A/duck/Nanjing/1102/2010(H4N8)	96.2	KC683700.1	A/duck/Nanjing/1102/2010(H4N8)	99.5	KC683700.1
PB1	A/avian/Egypt/920431/2006(H9N2)	97.8	GU050308.1	A/avian/Egypt/920431/2006(H9N2)	98.6	GU050308.1
PA	A/wild bird/Korea/A331/2009(H7N7)	99.1	JN244165.1	A/wild bird/Korea/A331/2009(H7N7)	99.4	JN244165.1
M	A/mallard/Sweden/113/02(H2N1)	99.7	CY122060.1	A/wild duck/Korea/CSM4-28/2010(H4N6)	99.5	JX454691.1
NP	A/chicken/Italy/9097/1997(H5N9)	98	EF597341.1	A/mallard/Korea/NHG187/2008(H7N7)	99.1	KC609905.1
NS	A/mallard/Bavaria/185-26/2008(H1N1)	98.8	HQ259236.1	A/mallard/ZhaLong/88/2004(H4N6)	98.7	FJ349250.1

TABLE 3 | Animal experiments with two H7N1 AIVs in chickens and ducks.

Birds	Viruses	Infected method	3 dpi		5 dpi		7 dpi		Sera-conversion (positive/total)	
			OP	CL	OP	CL	OP	CL	14 dpi	21 dpi
Chickens	BP/HuN/414/10	Inoculated	0/8[a]	0/8	0/5	0/5	0/5	0/5	5/5(2.19)[c]	5/5(4.15)
	(H7N1)	Contacted	0/3	0/3	0/3	0/3	0/3	0/3	0/3	2/3(1.15)
	CP/XH/420/10	Inoculated	6/8(<1)[b]	2/8(<1)	4/5(<1)	1/5(<1)	1/5(<1)	1/5(<1)	3/5(5.10)	4/5(5.67)
	(H7N1)	Contacted	0/3	0/3	1/3(<1)	0/3	0/3	2/3(<1)	1/3(4.69)	0/3
Ducks	BP/HuN/414/10	Inoculated	0/8	0/8	0/5	0/5	0/5	0/5	5/5(3)	0/5
	(H7N1)	Contacted	0/3	0/3	0/3	0/3	0/3	0/3	1/3(1.15)	0/3
	CP/XH/420/10	Inoculated	2/8(<1)	6/8(<1)	0/5	5/5(<1)	2/5(<1)	3/5(<1)	5/5(4.43)	5/5(5.51)
	(H7N1)	Contacted	1/3(<1)	0/3	0/3	3/3(<1)	0/3	2/3(<1)	3/3(2.31)	3/3(3.83)

Abbreviations: dpi, days post-inoculation; OP, oropharyngeal swab; CL, cloacal swab.

[a] *Positive birds/Total survival birds.*

[b] *Average viral titer of infected birds (log10EID50 ± SD).*

[c] *Average antibody titer of infected birds (log2).*

BP/HuN/414/10 in the oropharyngeal or cloacal samples from the chickens or ducks after inoculation (**Table 3**). However, the virus replicated at the tested organs of chickens, including kidney and bursal samples, and in the trachea of ducks (**Figures 2A,B**). In the inoculated group, seroconversion against BP/HuN/414/10 was detected by the HI assay in all chickens at 14 and 21 dpi (**Table 3**). In the contact group, seroconversion against BP/HuN/414/10 was detected by the HI assay in two of the three chickens at 21 dpi. However, the replication capacity of the isolated CP/XH/420/10 (H7N1) virus in ducks and chickens showed obvious differences. In the inoculated group, the virus was detected in chicken cloacal and oropharyngeal samples at 3, 5, and 7 dpi, but in the contact group, the virus was only detected at 5 dpi in one oropharyngeal sample and at 7 dpi in two of the cloacal samples. In the inoculated group, we detected the virus in duck cloacal samples at 3, 5, and 7 dpi, but in the oropharyngeal samples, the virus was only detected at 3 and 7 dpi. In the contact group, a virus titer was detected in one oropharyngeal sample at 3 dpi and in cloacal samples at 5 and 7 dpi (**Table 3**). At 3 dpi, viral replication was detected in many organs of the euthanized chickens and ducks (**Figures 2A,B**). To understand the antibody responses after infection, serum samples were collected from each bird for detection of antibodies by HI assay at 14 and 21 dpi: 60% of the inoculated chickens were seropositive at 14 dpi, and 80% were seropositive at 21 dpi. Whereas only 1 of the 3 contact chickens was seropositive at 14 dpi and all 3 were seronegative at 21 dpi. However, seroconversion was detected in all of the remaining ducks, suggesting that the isolated H7N1 CP/XH/420/10 strain stimulates a better immune response in ducks (**Table 3**).

Studies with Mice

To determine the capacity of the H7N1 AIVs to replicate and become pathogenic in mammals, we tested the two viruses in mice. The results showed that the two H7N1 isolates replicated in the lung, but the levels of the CP/XH/420/10 strain were higher than those of the BP/HuN/414/10 strain. In addition, the BP/HuN/414/10 strain replicated in the brain, and the CP/XH/420/10 strain replicated in the kidney and turbinate (**Figure 3A**). The difference in body weight loss at different dpi between the CP/XH/420/10 and controls groups was statistically significant ($P < 0.05$), but the difference between the BP/HuN/414/10 and controls groups was not ($P > 0.05$; **Figure 3B**). The pathological sections revealed pulmonary congestion and moderate broadening of the alveolar diaphragm due to the BP/HuN/414/10 (H7N1) virus infection in the mouse lung. CP/XH/420/10 (H7N1) caused pathological changes in many organs, including the lung, kidney and liver (**Figure 4**).

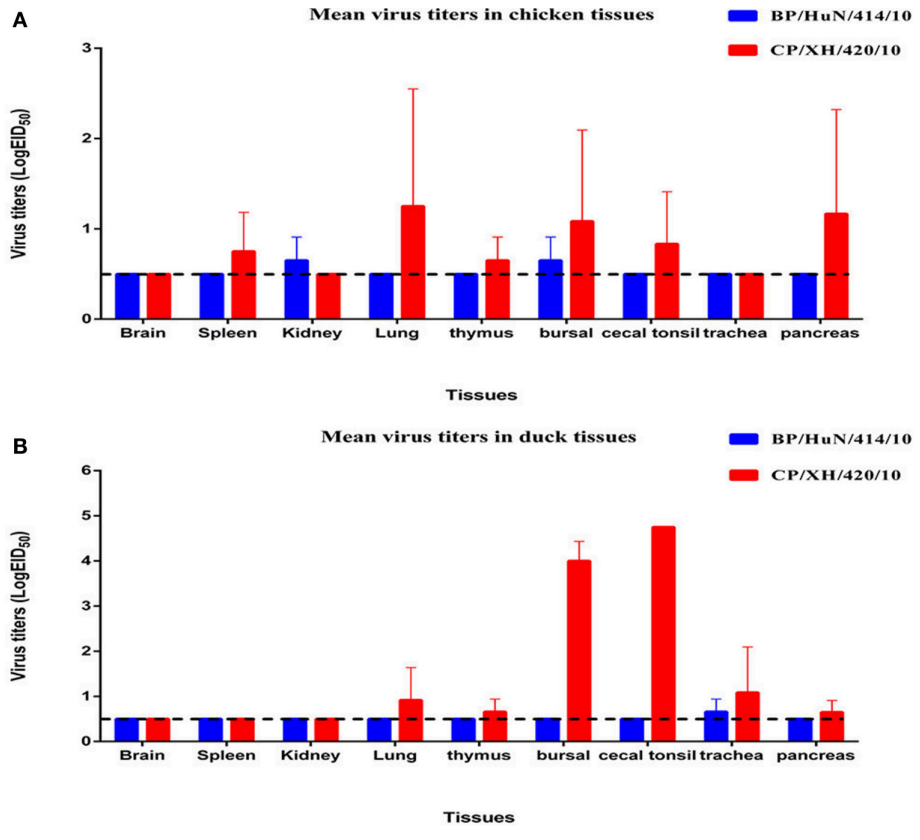


FIGURE 2 | Viral titers of the two isolated H7N1 viruses in the organs of chickens (A) and ducks (B) at 3 dpi. Viral titers are shown as the mean ± standard deviation. The dashed line indicates the lower limit of viral detection.

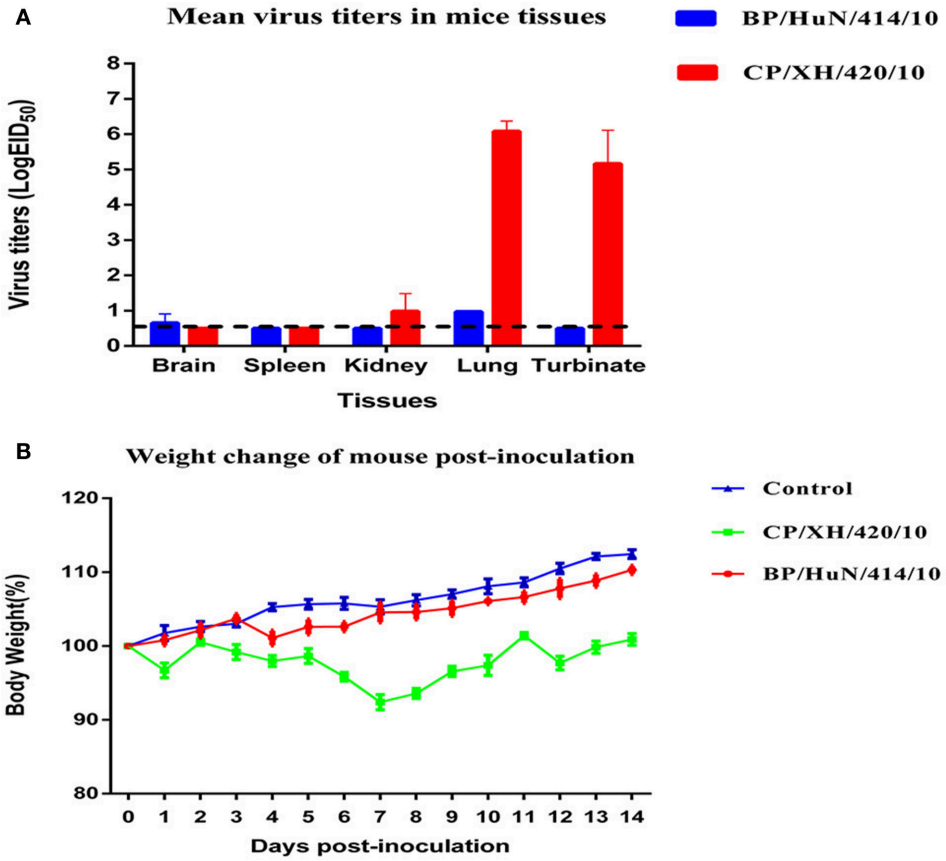


FIGURE 3 | Replication and virulence of the two isolated H7N1 viruses in mice: (A) Viral titers in the various tissues of the mice at 3 dpi. Viral titers are shown as the mean ± standard deviation. The dashed line indicates the lower limit of viral detection. **(B)** Weight changes in mice throughout the experiment. Weight changes are shown as the mean ± standard deviation. The weight changes of the CP/XH/420/10 and control groups differed significantly ($P < 0.05$). The BP/HuN/414/10 and control groups showed no significant difference ($P > 0.05$).

Analysis of Receptor Specificity of Two Strains of H7N1

The chicken red blood cells exhibit α-2,3-linked and α-2,6-linked sialic acid receptors. In contrast, the sheep red blood cells exhibited only α-2,3-linked sialic acid receptors.

The results showed that the A/Chicken/Jilin/HU/02 (H5N1) strain agglutinated chicken and sheep red blood cells but could not agglutinate chicken red blood cells treated with α-2,3-sialidase that have only α-2,6-linked sialic acid receptors indicating the avian receptor specificity. The A/Jilin/31/2005 (H1N1) strain agglutinated chicken red blood cells and chicken red blood cells treated with α-2,3-sialidase that have only α-2,6-linked sialic acid receptors but could not agglutinate sheep red blood cells that only have α-2,3-linked sialic acid receptors indicating the human receptor specificity. Furthermore, the two H7N1 AIVs agglutinated chicken red blood cells, chicken red blood cells treated with α-2,3-sialidase and sheep red blood cells, showing that they possess both avian and human receptor specificity (**Figure 5**).

DISCUSSION

LPAI H7 viruses, such as the H7N9 virus, were first reported in China in February and March 2013 and pose a serious threat to human health (Gao et al., 2013). The two isolated H7N1 viruses were used to study the phylogenetic and pathogenic characteristics of H7 LPAI's in this work. The two isolated H7N1 viruses both contain the motif PELPKGR↓GLFGAI at the cleavage site between HA1 and HA2, and the A/Shanghai/13/2013 (H7N9) strain carries the motif PEIPKGR↓GLFGAI at the cleavage site between HA1 and HA2 (**Table 1**); however, as none of these viruses shows serial basic amino acids in the motif, they meet the criteria for low pathogenicity (Steinhauer, 1999). In the analysis of the key amino acid sites, Q226L and G228S mutations were not found in the HA protein (**Table 1**), showing that the two H7N1 viruses retain the ability to preferentially bind to α-2,3-linked sialic acid receptors, which is a primary characteristic of AIVs (Yamada et al., 2006). Q226L and G228S mutations in the HA protein of A/Shanghai/13/2013 (H7N9) were not found. However,

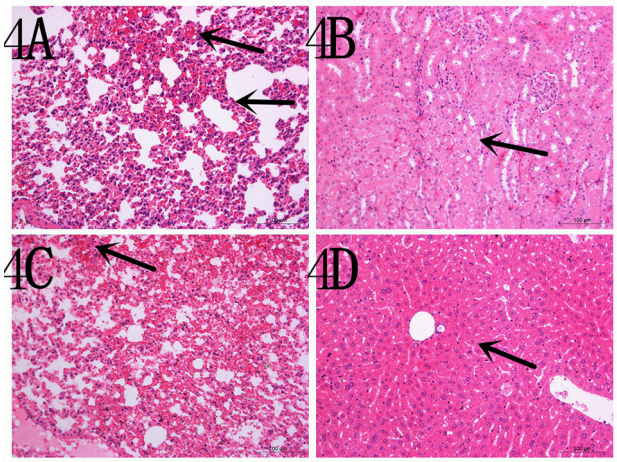


FIGURE 4 | Representative histopathological changes in the organs of the experimental mice (A) inoculated with the BP/HuN/414/10 (H7N1) virus and (B–D) with the CP/XH/420/10 (H7N1) virus. (A) Pulmonary congestion and moderate broadening of the alveolar septum. **(B)** Granular degeneration of renal tubular epithelial cells. **(C)** Pulmonary congestion. **(D)** Mild granular degeneration of some liver cells.

A/Shanghai/13/2013 (H7N9) and the two isolated H7N1 viruses presented mutations at positions S138A and T160A of the HA protein, which may favor mammalian adaptation and increase the affinity for α-2,6-linked sialic acid receptors (Ha et al., 2001; Wan and Perez, 2007; Fan et al., 2015). The two isolated viruses had no mutations at positions E627K and D701N of the PB2 protein, whereas A/Shanghai/13/2013 (H7N9) exhibits a mutation at position E627K, which plays an important role in the adaptation of AIVs to mammals (Katz et al., 2000; Li et al., 2005). No mutation at position S31N in the M2 protein was found, indicating that the two H7N1 viruses are sensitive to amantadine inhibitors (Lee et al., 2008), but a mutation was detected in the A/Shanghai/13/2013 (H7N9) strain. In addition, the two isolated H7N1 viruses and the A/Shanghai/13/2013 (H7N9) strain present mutations at position P42S of the NS1 protein, which may increase virulence in mice (Jiao et al., 2008). These results suggest that the two H7N1 AIVs can infect mammals and may have potential to infect humans.

Phylogenetic analysis based on the complete nucleotide sequences of the HA and NA genes showed that the two isolated LPAI H7N1 viruses cluster in the Eurasian branch. Homology analysis showed 98.4% homology between the entire genomes of the two H7N1 virus strains and the homologies of the eight gene segments of the two virus strains are 96.2% (PB2), 98.7% (PB1), 99.5% (PA), 99.3% (H7), 97.1% (NP), 99.3% (N1), 98.3% (M), and 99.8% (NS), indicating that both viruses most likely obtained their HA, NA, PA, and NS genes from a recent common ancestor, whereas they most likely obtained their other internal genes through reassortment with other influenza viruses. Given that the two strains originated from different places, with one from Hunan and the other from Jilin, the migration of birds may have played a role in promoting the spread of the virus over long distances.

No virus was detected in the oropharyngeal or cloacal samples of chickens or ducks after inoculation with the BP/HuN/414/10 virus. Chickens and ducks in the contact group also did not exhibit any viral titers, but seroconversion was detected in this group, suggesting that the virus can be transmitted within species by contact (**Table 3**). We also found an unusual phenomenon in that all 5 inoculated ducks were seropositive at 14 dpi, but no ducks were seropositive at 21 dpi. Similarly, it was odd that 1 of the 3 contact ducks exposed to the same virus (BP/HuN/414/10) was seropositive at 14 dpi, but none of the contact ducks were seropositive at 21 dpi (**Table 3**). These results show that the immunogenicity of the BP/HuN/414/10 virus is weak, and it is not able to stimulate the duck body to produce an effective humoral immune response. Therefore, the duration of the antibodies in the duck body is less than 21 days. The LPAI BP/HuN/414/10 virus replicated at the chicken kidney and bursal tissue as well as the duck trachea indicating that the virus can infect chickens and ducks. However, in contrast with BP/HuN/414/10, the replication and infection ability of the isolated CP/XH/420/10 (H7N1) virus in chickens and ducks showed obvious differences. Unlike BP/HuN/414/10, the virus CP/XH/420/10 (H7N1) was detected in both the oropharyngeal and cloacal samples of inoculated and contact chickens and ducks (**Table 3**), indicating that the isolated CP/XH/420/10 strain can be transmitted between species by contact. The HI assay showed that 60% of the chickens inoculated with CP/XH/420/10 were seropositive at 14 dpi, and 80% were seropositive at 21 dpi. However, only one of the 3 contacted chickens exposed to CP/XH/420/10 was seropositive at day 14, and none were seropositive at day 21 (**Table 3**). This difference in seropositivity between inoculated and contact chickens may be related to the amount of virus each chicken received as presumably the dose received by chickens in the directly inoculated group was higher than those in the contact group. In addition, the virus CP/XH/420/10 can replicate efficiently in multiple organs of chickens and ducks (**Figures 2A,B**), particularly in duck bursa and tonsils showing stronger replication capacity than BP/HuN/414/10. Ducks showed higher virus titer in bursa and tonsils which were avian immune organs is possible related to the immune system capture, aggregation and antigen clearing. Although, the pattern of infection in the two strains differed in some ways, both strains could infect chickens and ducks. Live poultry markets appear to be an important source of human infection (Chen et al., 2013; Li et al., 2014), and according to the literature, H7 subtypes of LPAI viruses have been reported to evolve into a highly pathogenic avian influenza virus in poultry (Horimoto and Kawaoka, 1995; Banks et al., 2001; Lee et al., 2005). Therefore, it is necessary to carry out AIV surveillance in live poultry markets.

In our study, the average weight change of mice inoculated with the two strains of H7N1 was slightly lower than that of the control group, suggesting that the two strains exhibit low pathogenicity in mice. The virulence of AIV in mice is determined by many amino acid sites, including PB2, HA, NS1, PA, and M1 mutations. The two isolated H7N1 viruses contained HA (138 and 160) and NS1 (42) mutations (**Table 1**). The virulence of CP/XH/420/10 was higher than

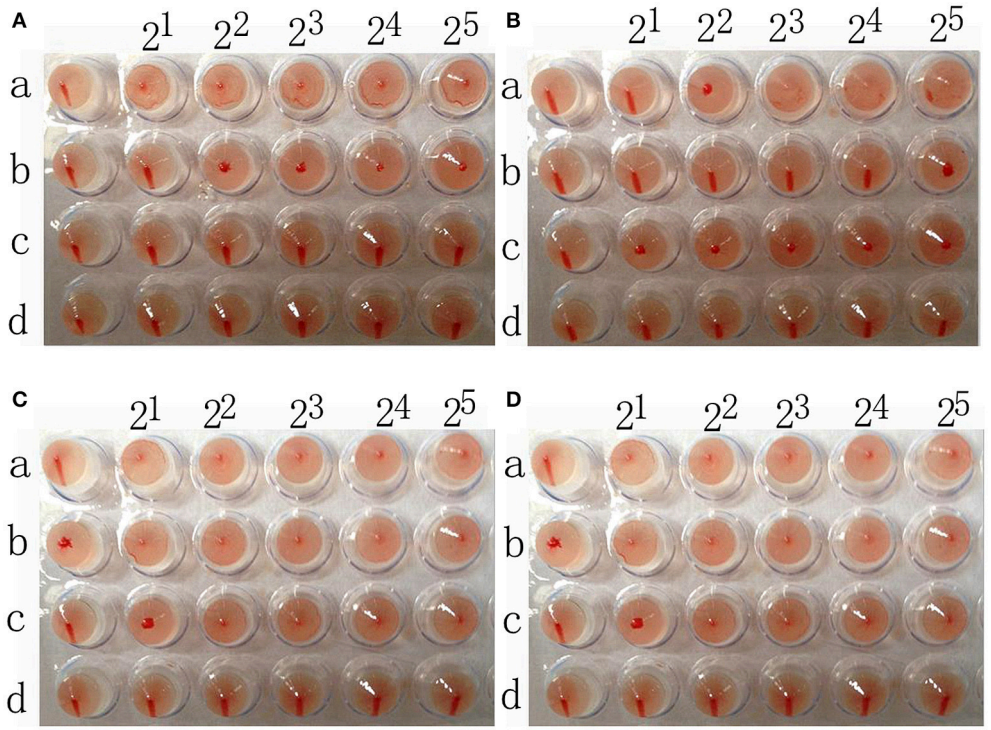


FIGURE 5 | The results of A/chicken/Jilin/HU/02 (H5N1) (A), A/Jilin/31/2005 (H1N1) (B), A/Baer's Pochard/Hunan/414/2010 (H7N1) (C), and A/Common Pochard/Xianghai/420/2010 (H7N1) (D) agglutination of different red blood cells. a: Chicken red blood cells (with α-2,3-linked sialic acid receptors and α-2,6-linked sialic acid receptors). b: Sheep red blood cells (with only α-2,3-linked sialic acid receptors). c: Chicken red blood cells treated with α-2,3-sialidase (with only α-2,6-linked sialic acid receptors). d: Chicken red blood cells treated with VCNA (no receptors).

TABLE 4 | Analysis of different amino acid sites in BP/HuN/414/10 (H7N1) and CP/XH/420/10 (H7N1).

Virus	HA (H7 numbering)			NA		PB2			NP	PB1		M1
	95	228	319	43	355	292	340	513	105	428	596	92
BP/HuN/414/10(H7N1)	I	A	K	R	D	V	R	F	M	H	S	N
CP/XH/420/10(H7N1)	V	T	R	Q	N	I	K	L	V	Q	P	K

that of BP/HuN/414/10, suggesting that some unknown factors may play important roles in the pathogenicity in mice (**Figure 3B**). According to the results of the animal experiments as a whole, including the infection of chickens, ducks and mice, the CP/XH/420/10 virus was more pathogenic than the BP/HuN/414/10 virus. We speculate that this phenomenon may be related to the differences in amino acid sites between the two viruses (**Table 4**), but we do not know which amino acid site would cause this difference. Therefore, these different amino acid sites will be studied by reverse genetics in our future work.

The HA protein's specificity for binding surface receptors in the host cell is a prerequisite for viral infection of the host. Human influenza viruses preferentially bind to α-2,6-linked sialic acid receptors, and avian influenza viruses preferentially bind to α-2,3-linked sialic acid receptors. In this study, the two isolated H7N1 viruses were able to bind to both α-2,3-linked sialic acid receptors and α-2,6-linked sialic acid receptors. This conclusion is consistent with the results that showed that BP/HuN/414/10 and CP/XH/420/10 could infect chickens, ducks and mice.

In conclusion, the results indicated that the two isolated H7N1 strains could infect mammals. Moreover, H7N1 AIVs are circulating widely in many areas, including North America (Panigrahy et al., 1995), Europe (Brown, 2010; Gonzales et al., 2011), southern China (Peng et al., 2014), New Zealand and Australia (Bulach et al., 2010). A previous study also found that an H7N1 AIV with no history of human infection could obtain capacity of airborne transmission among ferrets introduced by serial passage test (Sutton et al., 2014). Therefore, it is necessary to strengthen AIV surveillance in wild birds and poultry.

AUTHOR CONTRIBUTIONS

HJ contribute on experiment, writing article. DW contribute on isolation and identification of virus. JS collected the samples

and performed the experiment. HJ, DW, and JS are co-senior authors. YH, HC, YG, and YL guided the experiment and are the corresponding authors. YC contribute on sequencing. GC collected the samples. XZ, JZ, and XL helped to perform the experiments.

ACKNOWLEDGMENTS

This research was supported by the Fundamental Research Funds for Central Universities (2572014CA07), the Special Fund for Scientific Research with Public Interest in Forestry. In particular, this work was also funded by Surveillance of Wildlife Diseases from the State Forestry Administration of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01066

REFERENCES

Alexander, D. J. (2000). A review of avian influenza in different bird species. *Vet. Microbiol.* 74, 3–13. doi: 10.1016/S0378-1135(00)00160-7

Banks, J., Speidel, E., Moore, E., Plowright, L., Piccirillo, A., Capua, I., et al. (2001). Changes in the haemagglutinin and the neuraminidase genes prior to the emergence of highly pathogenic H7N1 avian influenza viruses in Italy. *Arch. Virol.* 146, 963–973. doi: 10.1007/s007050170128

Brown, I. H. (2010). Summary of avian influenza activity in Europe, Asia, and Africa, 2006-2009. *Avian Dis.* 54, 187–193. doi: 10.1637/8949-053109-Reg.1

Bulach, D., Halpin, R., Spiro, D., Pomeroy, L., Janies, D., and Boyle, D. B. (2010). Molecular analysis of H7 avian influenza viruses from Australia and New Zealand: genetic diversity and relationships from 1976 to 2007. *J. Virol.* 84, 9957–9966. doi: 10.1128/JVI.00930-10

Capua, I., Marangon, S., Dalla Pozza, M., Terregino, C., and Cattoli, G. (2003). Avian influenza in Italy 1997-2001. *Avian Dis.* 47, 839–843. doi: 10.1637/0005-2086-47.s3.839

Chen, H., Deng, G., Li, Z., Tian, G., Li, Y., Jiao, P., et al. (2004). The evolution of H5N1 influenza viruses in ducks in southern China. *Proc. Natl Acad. Sci. U.S.A.* 101, 10452–10457. doi: 10.1073/pnas.0403212101

Chen, Y., Liang, W., Yang, S., Wu, N., Gao, H., Sheng, J., et al. (2013). Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. *Lancet* 381, 1916–1925. doi: 10.1016/S0140-6736(13)60903-4

Fan, Z., Ci, Y., Liu, L., Ma, Y., Jia, Y., Wang, D., et al. (2015). Phylogenetic and pathogenic analyses of three H5N1 avian influenza viruses (clade 2.3. 2.1) isolated from wild birds in Northeast China. *Infect. Genet. Evol.* 29, 138–145. doi: 10.1016/j.meegid.2014.11.020

Fouchier, R. A., Schneeberger, P. M., Rozendaal, F. W., Broekman, J. M., Kemink, S. A., Munster, V., et al. (2004). Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc. Natl Acad. Sci. U.S.A.* 101, 1356–1361. doi: 10.1073/pnas.0308352100

Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., et al. (2013). Human infection with a novel avian-origin influenza A (H7N9) virus. *N. Engl. J. Med.* 368, 1888–1897. doi: 10.1056/NEJMoa1304459

Gonzales, J., van der Goot, J. A., Stegeman, J. A., Elbers, A. R., and Koch, G. (2011). Transmission between chickens of an H7N1 low pathogenic avian influenza virus isolated during the epidemic of 1999 in Italy. *Vet. Microbiol.* 152, 187–190. doi: 10.1016/j.vetmic.2011.04.022

Ha, Y., Stevens, D. J., Skehel, J. J., and Wiley, D. C. (2001). X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11181–11186. doi: 10.1073/pnas.201401198

Horimoto, T., and Kawaoka, Y. (1995). Molecular changes in virulent mutants arising from avirulent avian influenza viruses during replication in 14-day-old embryonated eggs. *Virology* 206, 755–759. doi: 10.1016/S.0042-6822(95)80004-2

Jiao, P., Tian, G., Li, Y., Deng, G., Jiang, Y., Liu, C., et al. (2008). A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. *J. Virol.* 82, 1146–1154. doi: 10.1128/JVI.01698-07

Katz, J. M., Lu, X., Tumpey, T. M., Smith, C. B., Shaw, M. W., and Subbarao, K. (2000). Molecular correlates of influenza A H5N1 virus pathogenesis in mice. *J. Virol.* 74, 10807–10810. doi: 10.1128/JVI.74.22.10807-10810.2000

Lee, C.-W., Swayne, D. E., Linares, J. A., Senne, D. A., and Suarez, D. L. (2005). H5N2 avian influenza outbreak in Texas in 2004: the first highly pathogenic strain in the United States in 20 years? *J. Virol.* 79, 11412–11421. doi: 10.1128/JVI.79.17.11412-11421.2005

Lee, J., Song, Y. J., Park, J. H., Lee, J.-H., Baek, Y. H., Song, M.-S., et al. (2008). Emergence of amantadine-resistant H3N2 avian influenza A virus in South Korea. *J. Clin. Microbiol.* 46, 3788–3790. doi: 10.1128/JCM.01427-08

Li, Z., Chen, H., Jiao, P., Deng, G., Tian, G., Li, Y., et al. (2005). Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* 79, 12058–12064. doi: 10.1128/JVI.79.18.12058-12064.2005

Li, Q., Zhou, L., Zhou, M., Chen, Z., Li, F., Wu, H., et al. (2014). Epidemiology of human infections with avian influenza A(H7N9) virus in China. *N. Engl. J. Med.* 370, 520–532. doi: 10.1056/NEJMoa1304617

Mancinelli, L., Onori, M., Concato, C., Sorge, R., Chiavelli, S., Coltella, L., et al. (2016). Clinical features of children hospitalized with influenza A and B infections during the 2012-2013 influenza season in Italy. *BMC Infect. Dis.* 16:6. doi: 10.1186/s12879-015-1333-x

Panigrahy, B., Senne, D., and Pearson, J. (1995). Presence of avian influenza virus (AIV) subtypes H5N2 and H7N1 in emus (*Dromaius novaehollandiae*) and rheas (*Rhea americana*): virus isolation and serologic findings. *Avian Dis.* 39, 64–67. doi: 10.2307/1591983

Peng, J., Yang, H., Jiang, H., Lin, Y.-X., Lu, C. D., Xu, Y.-W., et al. (2014). The origin of novel avian influenza A (H7N9) and mutation dynamics for its human-to-human transmissible capacity. *PLoS ONE* 9:e93094. doi: 10.1371/journal.pone.0093094

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* 27, 493–497.

Senne, D., Panigrahy, B., Kawaoka, Y., Pearson, J., Süss, J., Lipkind, M., et al. (1996). Survey of the hemagglutinin (HA) cleavage site sequence of H5 and H7 avian influenza viruses: amino acid sequence at the HA cleavage site as a marker of pathogenicity potential. *Avian Dis.* 40, 425–437. doi: 10.2307/1592241

Steinhauer, D. A. (1999). Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology* 258, 1–20. doi: 10.1006/viro.1999.9716

Stevens, J., Blixt, O., Tumpey, T. M., Taubenberger, J. K., Paulson, J. C., and Wilson, I. A. (2006). Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312, 404–410. doi: 10.1126/science.1124513

Sun, P., Xia, X., Hou, X., Gao, Y., He, W., Sun, H., et al. (2008). Development of a method for identification of receptor binding specificity of influenza virus. *Chin. J. Biol.* 21, 713–716. doi: 10.13200/j.cjb.2008.08.78.sunpl.016

Sutton, T. C., Finch, C., Shao, H., Angel, M., Chen, H., Capua, I., et al. (2014). Airborne transmission of highly pathogenic H7N1 influenza virus in ferrets. *J. Virol.* 88, 6623–6635. doi: 10.1128/JVI.02765-13

Wan, H., and Perez, D. R. (2007). Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. *J. Virol.* 81, 5181–5191. doi: 10.1128/JVI.02827-06

Yamada, S., Suzuki, Y., Suzuki, T., Le, M. Q., Nidom, C. A., Sakai-Tagawa, Y., et al. (2006). Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* 444, 378–382. doi: 10.1038/nature05264

Ye, G., Liang, C. H., Hua, D. G., Song, L. Y., Xiang, Y. G., Guang, C., et al. (2016). Phylogenetic analysis and pathogenicity assessment of two strains

of avian influenza virus subtype H9N2 isolated from migratory birds: high homology of internal genes with human H10N8 virus. *Front. Microbiol.* 7:57. doi: 10.3389/fmicb.2016.00057

Yu, Z., Sun, W., Li, X., Chen, Q., Chai, H., Gao, X., et al. (2015). Adaptive amino acid substitutions enhance the virulence of a reassortant H7N1 avian influenza virus isolated from wild waterfowl in mice. *Virology* 476, 233–239. doi: 10.1016/j.virol.2014.11.031

Zhao, D., Liang, L., Li, Y., Jiang, Y., Liu, L., and Chen, H. (2012). Phylogenetic and pathogenic analyses of avian influenza A H5N1 viruses isolated from poultry in Vietnam. *PLoS ONE* 7:e50959. doi: 10.1371/journal.pone.0050959

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Jin, Wang, Sun, Cui, Chen, Zhang, Zhang, Li, Chai, Gao, Li and Hua. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 November 2016
doi: 10.3389/fmicb.2016.01737

Two Genetically Similar H9N2 Influenza A Viruses Show Different Pathogenicity in Mice

*Qingtao Liu[1,2], Yuzhuo Liu[1,2], Jing Yang[1,2], Xinmei Huang[1,2], Kaikai Han[1,2], Dongmin Zhao[1,2], Keran Bi[1,2] and Yin Li[1,2]**

[1] Key Laboratory of Veterinary Biological Engineering and Technology, National Center for Engineering Research of Veterinary Bio-products, Institute of Veterinary Medicine, Ministry of Agriculture, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [2] Jiangsu Key Laboratory of Zoonosis, Jiangsu Co-Innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonoses, Yangzhou, China*

H9N2 Avian influenza virus has repeatedly infected humans and other mammals, which highlights the need to determine the pathogenicity and the corresponding mechanism of this virus for mammals. In this study, we found two H9N2 viruses with similar genetic background but with different pathogenicity in mice. The A/duck/Nanjing/06/2003 (NJ06) virus was highly pathogenic for mice, with a 50% mouse lethal dose (MLD_{50}) of $10^{2.83}$ 50% egg infectious dose (EID_{50}), whereas the A/duck/Nanjing/01/1999 (NJ01) virus was low pathogenic for mice, with a MLD_{50} of $>10^{6.81}$ EID_{50}. Further studies showed that the NJ06 virus grew faster and reached significantly higher titers than NJ01 *in vivo* and *in vitro*. Moreover, the NJ06 virus induced more severe lung lesions, and higher levels of inflammatory cellular infiltration and cytokine response in lungs than NJ01 did. However, only 12 different amino acid residues (HA-K157E, NA-A9T, NA-R435K, PB2-T149P, PB2-K627E, PB1-R187K, PA-L548M, PA-M550L, NP-G127E, NP-P277H, NP-D340N, NS1-D171N) were found between the two viruses, and all these residues except for NA-R435K were located in the known functional regions involved in interaction of viral proteins or between the virus and host factors. Summary, our results suggest that multiple amino acid differences may be responsible for the higher pathogenicity of the NJ06 virus for mice, resulting in lethal infection, enhanced viral replication, severe lung lesions, and excessive inflammatory cellular infiltration and cytokine response in lungs. These observations will be helpful for better understanding the pathogenic potential and the corresponding molecular basis of H9N2 viruses that might pose threats to human health in the future.

Keywords: H9N2, influenza A virus, genetic background, pathogenicity, mice

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Benjamin Roche,
French Research Institute for Development (IRD), France
Mariette Ducatez,
Institut National de la Recherche Agronomique, France

****Correspondence:***
Yin Li
muziyin08@163.com

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *08 August 2016*
Accepted: *17 October 2016*
Published: *04 November 2016*

Citation:
Liu Q, Liu Y, Yang J, Huang X, Han K, Zhao D, Bi K and Li Y (2016) Two Genetically Similar H9N2 Influenza A Viruses Show Different Pathogenicity in Mice. Front. Microbiol. 7:1737. doi: 10.3389/fmicb.2016.01737

INTRODUCTION

Avian influenza A viruses (AIVs) of the H9N2 subtype were first detected in turkeys in the United States in 1966 (Homme and Easterday, 1970), and have been circulating worldwide in multiple avian species and endemic in poultry populations across Eurasia (Alexander, 2000, 2007; Perk et al., 2006; Bi et al., 2010; Fusaro et al., 2011). Of note, H9N2 viruses in poultry have occasionally been transmitted to humans and other mammals (Peiris et al., 2001; Butt et al., 2005, 2010;

Sun et al., 2013), such as pigs and dogs, since this subtype was first reported to be detected in patients with influenza-like illness in Guangdong Province and in pigs in Hong Kong of China in 1998 (Guo et al., 1999; Lin et al., 2000; Peiris et al., 2001). In fact, follow-up serological surveys suggest that the incidence of human infections with H9N2 viruses might be more prevalent than what has been reported and possible human-to-human transmission cannot be completely excluded (Butt et al., 2005; Jia et al., 2008; Wang et al., 2009; Panwen et al., 2012). Clinically, human H9N2 infections present as typical seasonal influenza infections that can easily be overlooked (Lin et al., 2000; Butt et al., 2005), providing the viruses a greater opportunity to adapt to humans. These observations raise concerns about the possibility that H9N2 viruses might increase pathogenicity and transmissibility in humans. It is therefore important to investigate the pathogenicity and the corresponding mechanism of H9N2 viruses for mammals.

Previous studies revealed that some H9N2 viruses isolated from land-based poultry have demonstrated increased virulence for mammals. Guo et al. (2000) reported that the A/Chicken/Hong Kong/G9/97 and A/Quail/Hong Kong/G1/97 viruses could cause the deaths of three and two of the eight tested mice, respectively, at a dose of 10^6 50% egg infectious dose (EID_{50}) in mice, while Lu et al. reported that the two viruses are not lethal for mice at the same dose (Choi et al., 2004). In another study, Li et al analyzed 27 representative H9N2 viruses isolated from chickens and ducks in Mainland China, and found that some chicken isolates were able to replicate in mouse lungs efficiently and could induce a 10–20% weight loss of the inoculated mice, but none of the viruses are lethal for mice (Li et al., 2005). However, in 2007–2009, Bi et al. (2010) isolated six H9N2 viruses from chickens in northern China and found that these viruses could cause 50–85.7% mortality in mice at a dose of 10^6 EID_{50}. In addition, an H9N2 virus isolated from guinea fowl also showed enhanced replication and efficient transmission by direct contact in a ferret model (Wan et al., 2008). Although all these viruses cause death in mice at a high dose, none of the viruses are highly pathogenic for mice according to the criteria that a highly pathogenic virus has a 50% mouse lethal dose (MLD_{50}) value less than $10^{3.0}$ EID_{50} (Katz et al., 2000). However, several experimental evolutions by serial passage in mouse lungs showed that non-lethal H9N2 isolates could evolve to be lethal or highly pathogenic for mice after serial passage in mouse lungs (Zhang et al., 2011; Wang et al., 2012; Liu et al., 2014). Therefore, it is necessary to investigate the pathogenic potential and the corresponding molecular basis of H9N2 avian influenza viruses in mammals.

Several molecular determinants have already been identified that govern the pathogenicity of avian influenza virus for mammals, such as amino acid substitutions in the ribonucleoprotein (RNP) complex (Gabriel et al., 2005; Salomon et al., 2006; Song et al., 2009; Sun et al., 2015), the mutations involved in the ability of NS1 proteins to restrict the induction of the host interferon response (Li et al., 2006; Thulasi Raman and Zhou, 2016), the length of the NA stalk (Zhou et al., 2009). However, most of these studies focused on H5 subtype of influenza viruses, and the pathogenic mechanism of H9N2 viruses for mammals is poorly understood. The substitution PB2 E627K which has been shown to be a key factor in the increased virulence of H5N1 AIVs to mammals has also been observed in the adaption of H9N2 viruses in mice (Zhang et al., 2011; Wang et al., 2012). However, Wang et al. (2012) reported that although the E627K mutation on its own enhanced replication and polymerase activity, it did not significantly increase pathogenicity of H9N2 virus, and only the combination of PB2 E627K and M147L could increase the virulence of the H9N2 virus in mice. Another report also showed that a H9N2 virus containing a human-like PB2 segment with 627 K is non-pathogenic for mice, while the mutation F404L in the PB2 segment could increase the virulence of the H9N2 virus and the combination of PB2 F404L with mutations in PA (D3V and S225R) and HA (L80F and N193D) was able to make the non-pathogenic H9N2 virus become high pathogenic for mice (Liu et al., 2015). However, it is still unknown that if H9N2 virus could acquire the mutations that govern the high pathogenicity for mice in natural, especially in poultry.

In this study, we characterized two H9N2 AIVs, NJ06 and NJ01, that were isolated from ducks in China. The NJ06 virus was highly pathogenic for mice and induced severe lung lesions and excessive cytokine responses, while the NJ01virus exhibited low pathogenicity in this model. However, there were only twelve amino acid differences between the two viruses, which might contribute to the high virulence of the NJ06 virus in mice. Therefore, the two viruses had similar genetic background, but showed different pathogenicity for mice, which offer an appropriate system in which to explore the molecular basis of host adaptation and enhanced virulence in mammals.

MATERIALS AND METHODS

Ethics Statements

All animal experiments were approved by the Committee on the Ethics of Animal Experiments of Jiangsu Academy of Agricultural Sciences (JAAS no. 20141107), and complied with the guidelines of Jiangsu Province Animal Regulations (Government Decree No. 45). All experiments involving live viruses and animals were carried out in negative pressure isolators with HEPA filters in a biosafety level 2+ laboratory (enhanced animal biosafety level 2 laboratory and a negative pressure-ventilation laboratory) in accordance with the institutional biosafety manual.

Viruses and Cells

The H9N2 viruses A/duck/Nanjing/06/2003 (NJ06) and A/duck/Nanjing/01/1999 (NJ01) were isolated from ducks in Jiangsu, China and propagated in specific pathogen-free (SPF) embryonated chicken eggs. Viral titers were measured by calculating the EID_{50}. Madin–Darby canine kidney (MDCK) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% fetal bovine serum.

Mouse Studies

Female BALB/c mice (5 weeks old) were used in this study. To evaluate the virulence of the NJ06 and NJ01 viruses, groups of five BALB/c mice were anesthetized with pentobarbital natricum and inoculated intranasally with 10-fold serial dilutions of viruses in 30 μl PBS or mock inoculated with PBS to serve as controls. Body weight and survival of mice were recorded daily for 14 days. Mice that showed severe symptoms or lost more than 25% of their body weight were euthanized and scored as dead for humane reasons. The MLD_{50} of virus was calculated and expressed in EID_{50}.

To evaluate viral replication in mice, groups of female BALB/c mice were intranasally inoculated with 10^5 EID_{50} of the NJ01 and NJ06 viruses, respectively. At 1, 2, 3, and 5 days post inoculation (dpi), five mice in each group were euthanized, and whole lungs were removed and homogenized in 1 ml of PBS for virus titration in 10-day-old embryonated eggs as previously described (Hu et al., 2013).

To assess lung injury, groups of five BALB/c mice were intranasally inoculated with 10^5 EID_{50} of the NJ01 and NJ06 viruses, respectively, and lung histopathology and water content were determined at 5 dpi. For histopathological analysis, mouse lungs were fixed in 4% paraformaldehyde, embedded in paraffin, cut into 5 mm-thick sections and then stained with haematoxylin and eosin (H&E) for light microscopy. For water content analysis, mouse lungs were surgically dissected, blotted dry, and weighed immediately as wet weight, and then dried in an oven at 80°C for 72 h and reweighed as dry weight. The lung wet/dry weight ratios were calculated for each animal to assess tissue edema as previously described (Lang et al., 2005).

Growth Properties *In vitro*

To evaluate the replication of virus *in vitro*, MDCK or A549 cell monolayer in 12-well plates were washed three times with PBS, and inoculated at a multiplicity of infection of 0.01, overlaid with serum-free DMEM containing 2 mg/ml TPCK-trypsin (Sigma–Aldrich). Virus titers in supernatants were determined as the number of 50% tissue culture infectious doses ($TCID_{50}$) per ml in MDCK cells at 12, 24, 48, and 72 h post inoculation (hpi).

Differential Leukocyte Counts and Cytokine Expression Analysis

Inflammatory response in mouse lungs was assessed by testing differential leukocyte counts in bronchoalveolar lavage (BAL) fluid and expression profiles of representative cytokine genes of mice infected with 10^5 EID_{50} of the NJ06 or NJ01 virus at the indicated days. To determine differential leukocyte counts, BAL cells were obtained from mouse lungs in each group as described by Nick et al. (2000) and Densmore et al. (2013). In brief, the lungs were lavaged twice with a total 1 ml saline (4°C) through the endotracheal tube, and the recovery rate of BAL fluid was not less than 90% for each animal tested. After the amount of fluid recovered was recorded, an aliquot of BAL fluid was diluted 1:1 with 0.01% crystal violet dye and 2.7% acetic acid for leukocyte staining and erythrocyte hemolysis, and the number of leukocytes in BAL fluid was counted with a haemocytometer under a light microscope. Subsequently, the remaining fluid was centrifuged for 10 min at 300 × *g*. Cell differential counts were determined by Wright staining of a spun sample, on the basis of morphological criteria under a light microscope with evaluation of at least 200 cells per slide, and each slide was counted twice by different observers blinded to the status of the animal.

Quantitative real-time PCR (qRT-PCR) was used to analyze the expression of cytokine genes in mouse lungs. Total RNA was isolated from lungs using TRIzol reagent (Life Technologies) and treated with DNase I (Fermentas, Glen Burnie, MD, USA). One microgram of total RNA per sample was reverse transcribed into cDNA using a PrimeScript RT Reagent Kit (Takara). The cDNA was run in the ABI 7500 Real Time PCR System using an SYBR Premix Ex Taq Kit (Takara). One cycle for melting curve analysis for all reactions was added to verify product specificity. The expression of each cytokine gene relative to that of the β-actin was calculated using the $2^{-\Delta\Delta CT}$ method. The primers for TNF-α, CXCL10, IL-17a, and IL-10 were designed based on these target mouse genes with GenBank accession numbers NM_013693.3, NM_021274.2, NM_010552.3, and NM_010548.2, respectively. The primers for β-actin have been described previously (Hu et al., 2013). Primers for theses target genes were as follows (forward and reverse primers, respectively): for TNF-α, 5′-GCCAGGAGGGAGAACAGAAACTC-3′ and 5′-GGCCAGTGAGTGAAAGGGACA-3′; for CXCL10, 5′-ATC CGGAATCTAAGACCATCAAGAA-3′ and 5′-TGTCCATCCAT CGCAGCAC-3′; for IL-17a, 5′-GAAGG CCCTCAGACTA CCTCAA-3′ and 5′-TCATGTGGTGGTCCAGCTTTC-3′; for IL-10, 5′-GCCAGAGCCACATGCTCCTA-3′ and 5′-GATAA GGCTTGGCAACCCAAGTAA-3′; and for β-actin, 5′-CATCCG TAAAGACCTCTATGCCAAC-3′ and 5′-ATGGAGCCACCG ATCCACA-3′.

Sequence Analysis

Viral RNA was extracted from infected allantoic fluid using the Body Fluid Viral DNA/RNA Kit (Axygen) according to the manufacture's protocol. Reverse transcription was performed using the Uni12-primer (5′-AGCAAAAGCAGG-3′) by standard methods, and PCR amplification of cDNA was directed with previously described primers (Hoffmann et al., 2001). PCR products were purified using a DNA Gel Extraction Kit (Axygen) in accordance with manufacture's recommendations, and cloned into the pMD-18T vector (Takara), and sent for commercial sequence analysis (Sangon Biotechnology, Shanghai, China). Sequencing results were phylogenetically analyzed with the representative strains available in GenBank. The nucleotide sequences were initially aligned using the Clustal V alignment algorithm of the Megalign program (DNAStar, Madison, WI, USA). The phylogenetic tree was constructed using MEGA 6.06 software with the neighbor-joining method. The GenBank accession numbers for theNJ01 segments are KX349960, KX349961, DQ681205, and DQ681221 to DQ681225, and those for the NJ06 segments are KX349952 to KX349959.

Statistical Analysis

Data were analyzed using the SPSS Statistics software and results were expressed as means ± standard deviation (SD).

The statistical significance of differences was determined by an independent-sample *t*-test.

RESULTS

Pathogenicity of NJ01 and NJ06 Viruses in Mice

To compare the virulence of NJ06 with that of NJ01 virus, MLD_{50} was determined. The results showed that NJ06 owned an MLD_{50} of $10^{2.83}$ EID_{50} and was highly pathogenic for mice based on MLD_{50} values of $<10^{3.0}$ EID_{50} (Katz et al., 2000; Chen et al., 2004), while NJ01 was low pathogenicity in mice, with an MLD_{50} of $>10^{6.81}$ EID_{50}. The morbidity and mortality of the two viruses were also compared in mice in another study. NJ06-infected mice showed obvious signs of illness, including decreased activity, huddling, ruffled fur, heavy/labored breathing, and hunched posture. Mice in this group began to lose weight at 1 dpi with a high inoculated dose of $10^{6.0}$ EID_{50}, and at 3 dpi with a low inoculated dose of $10^{3.0}$ EID_{50} (**Figure 1A**). In addition, all of the mice in the NJ06-infected group showed obvious weight loss and died by 5 dpi at a dose of $10^{6.0}$ EID_{50}, and by 10 dpi at a dose of $10^{4.0}$ EID_{50} (**Figure 1B**). In contrast, no mortality was observed in NJ01-infected mice (**Figure 1D**), and the mice in this group displayed only slight weight reduction throughout the course of infection, and started to gain weight at 5 dpi, even at a high dose of $10^{6.0}$ EID_{50} (**Figure 1C**).

Replication of NJ06 and NJ01 Viruses *In vivo* and *In vitro*

To determine whether the differences in virulence of NI06 and NJ01 were related to the differences of viral replication in mice, the levels of viral replication in mouse organs were compared. No infectious virus was detected in the heart, liver, spleen, kidney, brain from any of the NJ06- or NJ01-infected mice, whereas both the two viruses replicated well in mouse lungs. However, the NJ06 virus replicated to a high titer that was $10^{1.4}$-fold higher than the NJ01 virus as early as 1 dpi and sustained significantly higher levels of replication than the NJ01 virus throughout the course of infection (**Figure 2A**). The NJ06 virus reached a peak titer of $10^{7.4}$ EID_{50}/ml at 3 dpi, versus a peak titer of $10^{4.9}$ EID_{50}/ml at this time point for NJ01. Therefore, the NJ06 virus grew faster and to significantly higher titers than the NJ01 virus, though the two viruses all could replicate in mouse lungs.

The replication kinetics of the two viruses *in vitro* was also measured in MDCK and A549 cells. The NJ06 virus grew to significantly higher titers than NJ01 in either MDCK or A549 cells at each time point (**Figure 2B**). The NJ06 virus reached a peak titer of $10^{7.8}$ $TCID_{50}$/ml at 48 hpi in MDCK cells, which was $10^{2.6}$-fold higher than the peak titer of NJ01 at this time

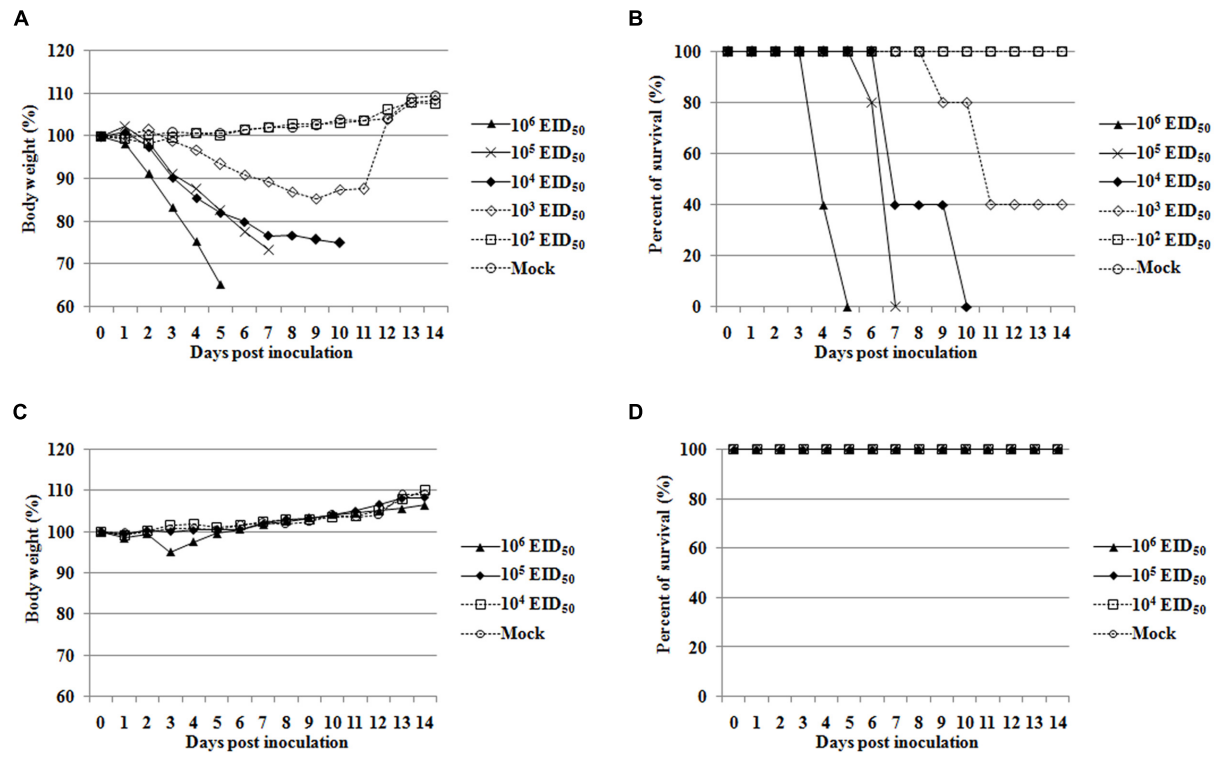


FIGURE 1 | Pathogenicity of NJ06 and NJ01 viruses in mice. Groups of five BALB/c mice were inoculated intranasally with different doses of virus (10^2–10^6 EID_{50} of NJ06 or 10^4–10^6 EID_{50} of NJ01) or mock inoculated (Mock). Morbidity was evaluated by monitoring weight changes over a 14-day period and is graphed as a percentage of the animals' weight on the day of inoculation (day 0). The average body weight of mice infected with NJ06 **(A)** or NJ01 **(C)** is shown. Mortality of NJ06 **(B)** and NJ01 **(D)** infected mice was examined by measuring the percent survival.

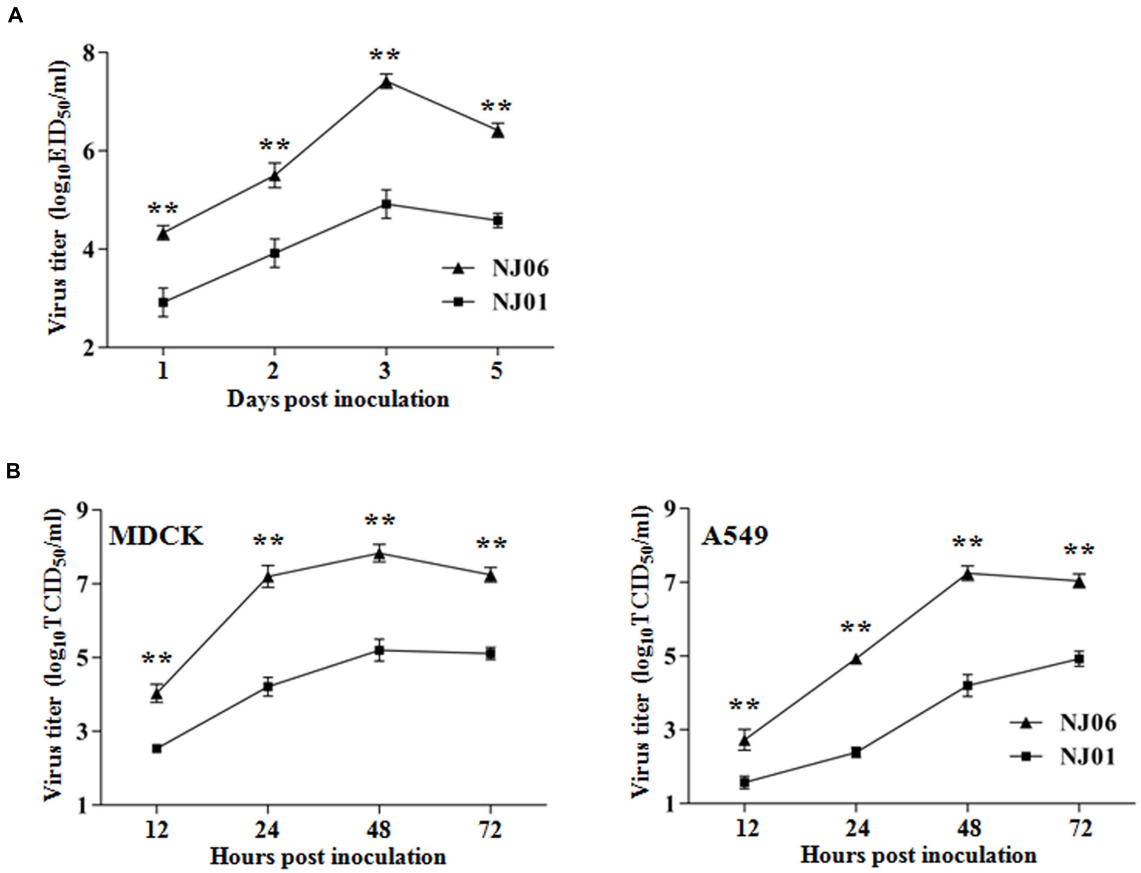


FIGURE 2 | Replication kinetics of NJ06 and NJ01viruses *in vivo* and *in vitro*. (A) Groups of three BALB/c mice were inoculated intranasally with 10^5 EID_{50} of NJ06 or NJ01 viruses, and lungs were collected at for virus titration in eggs on the indicated days. **(B)** Madin–Darby canine kidney (MDCK) or A549 cells were infected with NJ06 or NJ01 viruses at a multiplicity of infection of 0.01, and virus yields were determined in MDCK cells at appropriate time points. Data represent the means of the results determined for three independent experiments ± standard deviations. $^{**}P < 0.01$ compared with the value for the NJ01 virus.

point. The peak viral titer of NJ06 virus in A549 cells was also observed on 48 hpi, reaching $10^{7.2}$ $TCID_{50}$/ml, versus virus titer of $10^{4.2}$ $TCID_{50}$/ml at 48 hpi and reaching peak yield with $10^{4.9}$ $TCID_{50}$/ml at 72 hpi for NJ01. Therefore, the replication abilities of NJ06 virus were significantly higher than the NJ01 virus both *in vivo* and *in vitro*, which might correlate with the higher virulence of NJ06 virus in mice.

Severe Lung Lesions in Mice Infected with NJ06 Virus

To compare the lung lesions of mice infected with NJ06 or NJ01, the gross and histopathologic changes in the mouse lung was determined at 5 dpi, a time that immediately preceded the death of mice infected with the NJ06 virus. We found that NJ06-infected mice exhibited severe edema, congestion, and hemorrhage in lungs (**Figure 3A**, right), whereas the lung of NJ01-infected mice appeared normal except for occasional small dark red foci of pneumonia (**Figure 3A**, left). Moreover, the severe pulmonary edema in NJ06-infected mice was further confirmed by the lung wet/dry weight ratios (**Figure 3B**). Histologically, NJ06 induced severe pneumonia with inflammatory cellular infiltration and hemorrhage, alveolar wall edema and thickening, and deciduous epithelium mucosae and inflammatory cells in the bronchioles (**Figures 3C,D**). However, only mild and limited alveolitis was observed in the lungs of NJ01-infected mice (**Figures 3E,F**).

Increased Numbers of BAL Cells in Mice Infected with NJ06 Virus

To better characterize the inflammatory cellular components in lungs, total and differential cell counts in BAL fluid were determined at 5 dpi for NJ06- and NJ01-infected mice, respectively. The total number of BAL cells was markedly increased in mice infected with NJ06 virus compared to the NJ01 virus or mock infection group (**Figure 4A**). By contrast, there were no statistically significant differences in the total numbers of BAL cells between the NJ01 and mock infection groups. In addition, cellular infiltration in the BAL samples during NJ06 virus infection was associated with an increase in the percentage of neutrophils and lymphocytes compared with that for NJ01-infected mice (**Figures 4B–D**). These data suggested that NJ06 virus induced a much larger increase of inflammatory cell

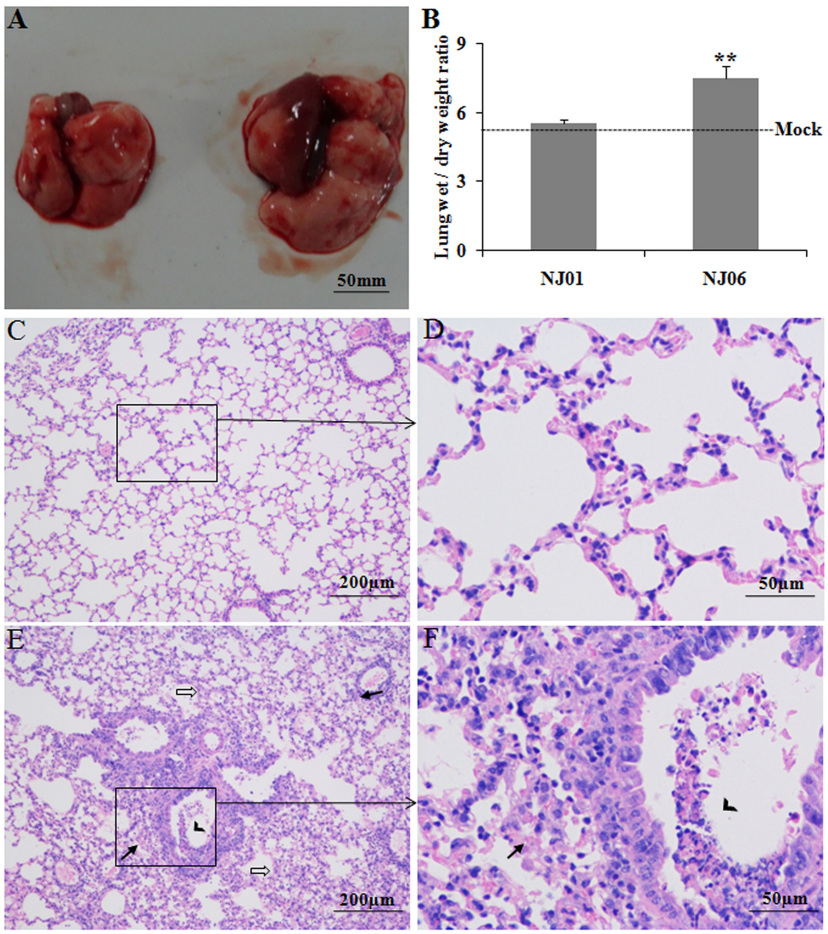


FIGURE 3 | Lung lesions in mice. Groups of BALB/c mice were intranasally inoculated with 10^5 EID_{50} of the NJ06 and NJ01 viruses, respectively, gross pathology, water content, and histopathology of lungs were determined at 5 days post inoculation (dpi). **(A)** Gross pathology of H9N2virus-infected lung. Severe edema, congestion, and hemorrhage were observed in NJ06-infected lung (left), and the NJ01-infected lung (right) appeared normal except for occasional small dark red foci. **(B)** Lung wet/dry weight ratios of H9N2 virus-infected mice. Data represent the means of the results determined for three independent experiments ± standard deviations. ** indicates $p < 0.01$ compared with the mock group. **(C,D)** Representative histopathological changes in NJ01-infected lungs (H&E-stained). Only mild and limited alveolitis was observed. **(E,F)** Representative histopathological changes in NJ06-infected lungs. Solid arrows indicate inflammatory cell infiltrates around the bronchus, open arrows indicate edema and thickening of alveolar walls, Solid arrowheads indicate desquamation of epithelial cells in bronchial lumens.

infiltrate into the lungs, especially neutrophils and lymphocytes, which may be contributing to the pathogenesis of the severe lung injury associated with NJ06 virus infection.

NJ06 Virus Elicits Significantly High Levels of Cytokine Response in Mouse Lungs

To determine whether the different levels of virulence of NJ06 and NJ01 viruses were related to difference in cytokine expression levels induced by the two viruses in mice, lungs from mice infected with 10^5 EID_{50} of NJ06 or NJ01 virus were collected, and subsequently assayed for TNF-α, CXCL10, IL-17a, and IL-10 levels by qRT-PCR, respectively. The levels of all cytokines were substantially greater than constitutive levels in the lungs of NJ06- or NJ01-infected mice by 1 dpi (**Figure 5**). However, NJ06 virus induced significantly higher levels of TNF-α and CXCL 10 expression than did NJ01 at all time points. The IL-17a was also detected at higher levels of expression in NJ06-infected mice than in NJ01-infected mice, although the result at the early time point, 1 dpi, was not significant. By contrast, at 1 dpi, the levels of IL-10 were reduced in NJ06-infected mice compared with those in NJ01-infected mice. Although the levels of IL-10 were elevated in NJ06-infected mice at 3 and 5 dpi, the latter result was not significant compared with levels found in NJ01-infected mice. In addition, a continuous increase in levels of IL-10 was observed throughout the entire study period in NJ01-infected mice, whereas the IL-10 levels in NJ06-infected mice were reduced at the end of the observational period (5 dpi) compared with the former time point (3 dpi). These data suggest that the induction of inflammatory cytokines by NJ06 is different from the induction by NJ06, which might contribute to the observed differences in severity of disease caused by the two viruses.

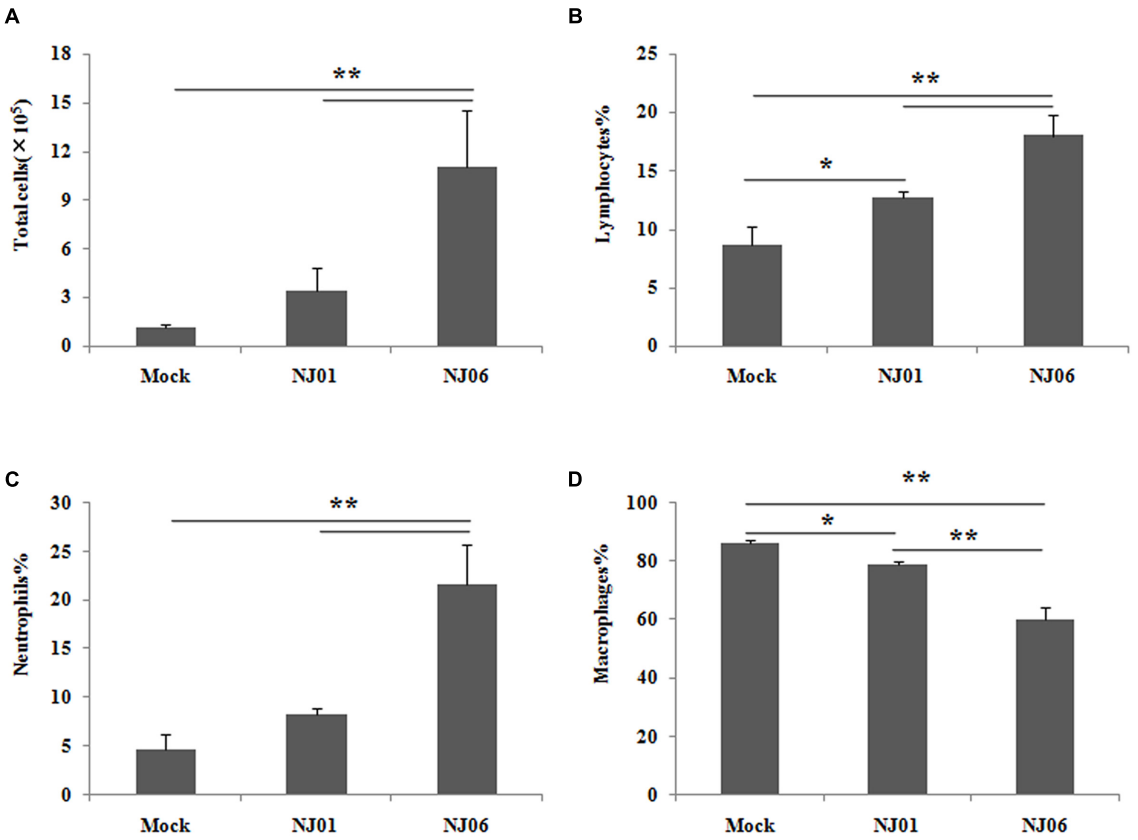


FIGURE 4 | Total and differential cell counts in bronchoalveolar lavage (BAL) fluid. Groups of five mice were inoculated intranasally with 10^5 EID_{50} of NJ06 or NJ01 virus, and BAL specimens were collected to determine cell counts on 5 dpi. The total numbers of BAL cells **(A)** and percentages of lymphocytes **(B)**, neutrophils, **(C)** and macrophages **(D)** of virus- or mock-inoculated mice are shown as means ± SD. *$p < 0.05$ and **$p < 0.01$.

Sequence and Phylogenetic Analysis

To determine the molecular basis for the differences in pathogenicity between the two viruses, the sequences of all of eight segments of NJ06 were compared with those of NJ01 virus. This revealed twelve amino acid differences between these two viruses, which were mapped to PB2, PB1, PA, HA (H3 numbering used throughout the text), NP, NA, and NS gene (**Table 1**). Phylogenetic analysis of the HA genes showed that both viruses belonged to the Ck/BJ/1/94-like lineage (**Figure 6**), with the same R-S-S-R amino acid motif at the cleavage sites, a characteristic of low pathogenic avian influenza virus (LPAIV) between HA1 and HA2 [4,34]. The NA and M genes of these two isolates also belong to the Ck/BJ/1/94-like lineage (Supplementary Figure S1), and both viruses had the same "marking" deletion of three amino acids (positions 62–64) at the NA stalk region, as previously described (9, 11, 15). The NS and ribonucleoprotein (PB2, PB1, PA, and NP) complex genes of the two viruses fell into the DK/HK/Y439/97-like and Ck/SH/F/98-like, respectively (Supplementary Figure S1).

To find out whether the amino acids found at these positions in NJ06 were present also in other natural H9N2 strains, we analyzed the H9N2 sequences deposited in the Influenza Research Database (the National Institute of Allergy and Infectious Diseases database[1]; **Table 1**). Most avian, swine, and human isolates possess the same amino acid as NJ06 at positions PB1-187 (R), HA-147 and -477 (K and Y), NP-277 and -340 (P and D), NA-9 and -435 (A and R), and NS1-171(D). By contrast, at positions PB2-149 and -627, PA-548 and -550, and NP-127, most avian, swine, and human isolates share common residues as NJ01 virus. In fact, only five and one avian isolates share the same residues with the NJ06 virus at positions PB2-627(K) and NP-127 (G), respectively, and no isolates was found to contain the residues PB2-149T, and PA-548L and -550M as NJ06 virus. Therefore, the amino acids observed at these positions in NJ06 were unique to this virus, which might contribute to the high virulence of the NJ06 virus.

DISCUSSION

Although highly pathogenic avian influenza viruses (HPAIVs), such as H5 and H7 viruses, have caused serious harm to human health, some recent studies have suggested that LPAIVs, especially H9N2 viruses, could jump to humans more easily (Wan

[1]http://www.fludb.org

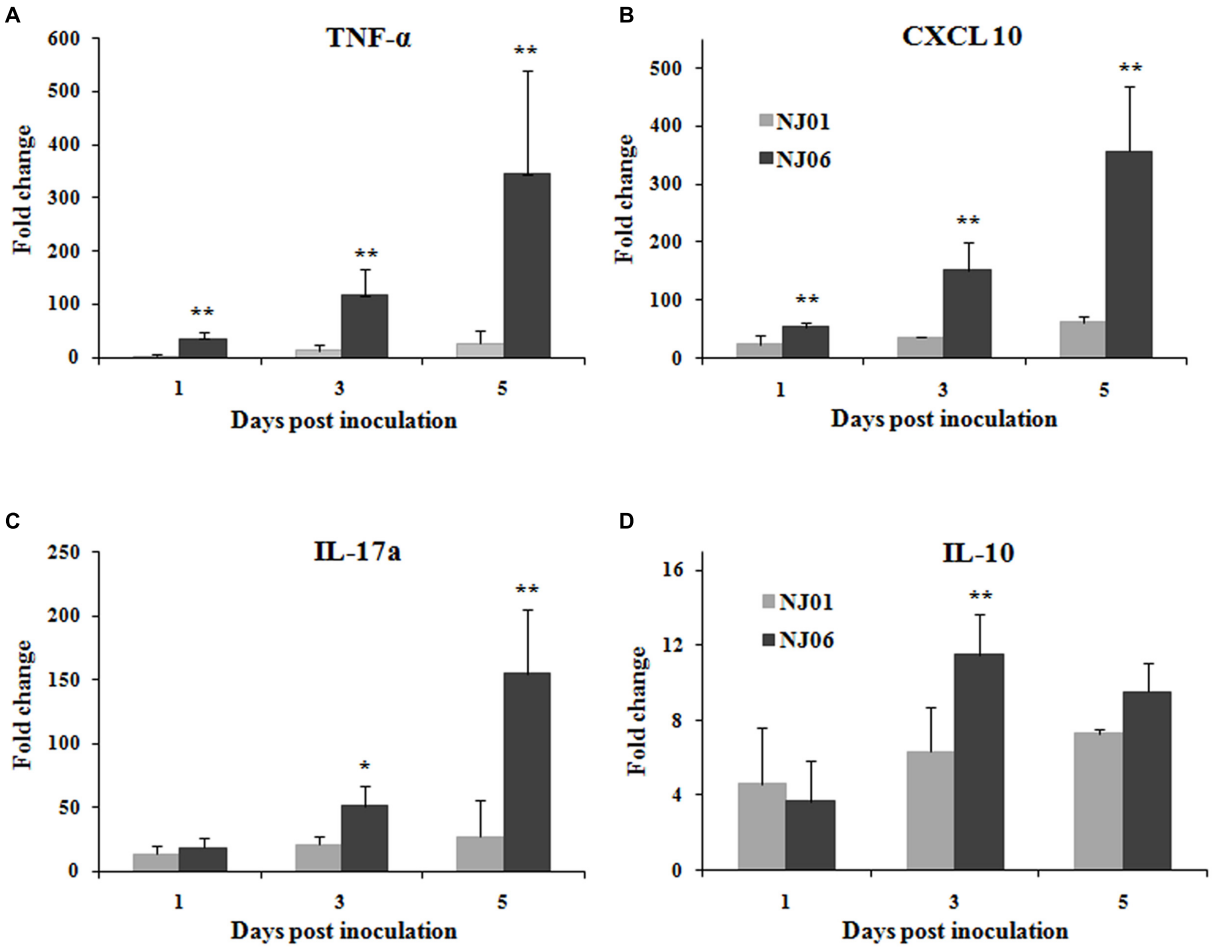


FIGURE 5 | Inflammatory cytokines response in mouse lungs. Groups of five mice were intranasally inoculated with 10^5 ELD_{50} of NJ06 or NJ01 virus, and the expression of these cytokines, TNF-α, IL-1β, IL-8, and IL-10, was measured in the mouse lungs at the indicated times by Quantitative real-time PCR (qRT-PCR). Each cytokine expression was normalized to the expression of β-actin, and is presented as the fold change relative to the level of the mock group. The mean fold change of each group is shown, with error bars representing the SD. * indicates $p < 0.05$ and ** indicates $p < 0.01$ compared with the NJ01 virus infection group.

and Perez, 2007; Long et al., 2015). H9N2 AIVs have repeatedly infected humans and other mammals (Peiris et al., 1999; Yu et al., 2011; Sun et al., 2013), such as pigs and dogs, and could cause mild respiratory disease in humans (Guo et al., 1999; Peiris et al., 1999). More seriously, some H9N2 AIVs isolates could replicate efficiently in mice and ferrets without prior adaptation, and was able to adapt to high pathogenicity in mice. All these facts indicate that H9N2 AIVs have gradually acquired mutations that make them more adapted to mammals including humans (Wan and Perez, 2007; Kimble et al., 2011; Mok et al., 2011), posing a significant threat to public health. Therefore, it is necessary to investigate the pathogenesis of H9N2 AIVs in mammals.

Although the ferret is well established as an animal model to study human influenza virus pathogenesis and transmission, no animal model is perfect and the use of ferrets for influenza studies has been limited by the lack of availability of inbred and specific pathogen–free animals, and the corresponding immunological reagents (Oh and Hurt, 2016). Therefore, the mouse, another commonly used model in influenza virus research (Belser and Tumpey, 2013; Cauldwell et al., 2014; Thangavel and Bouvier, 2014), was used to compared the pathogenicity of the two genetically similar H9N2 viruses in this study. We found that the NJ06 virus was highly pathogenic for mice, while the NJ01 virus exhibited low pathogenicity in this animal model. The NJ06 virus caused signs of severe disease and resulted in 60% mortality at a low inoculation dose of 10^3 EID_{50}, whereas infection with the DK1 virus did not cause death or obvious clinical signs of illness even at a high dose of 10^6 EID_{50}. Previous studies showed that the high virulence of H5N1 AIVs for mice is associated with the enhanced replication and extra-pulmonary infection (Sirinonthanawech et al., 2011; Hu et al., 2013). Here, the NJ06 virus grew faster and to significantly higher titers in mouse lungs than NJ01 virus, but both the two viruses were not able to spread to the extra-pulmonary organs. These data support the viewpoint that high replication ability replication in lungs is an important and characteristic prerequisite for high virulence of AIV in mice.

Severe lung lesions characterized by massive edema, diffuse alveolar damage, and excessive inflammatory cell infiltration

TABLE 1 | Amino acid differences between the NJ06 and NJ01 viruses.

Viruses	Amino acid residue at indicated position (Number of strains possessing mutation/total number of strains examined)													
	PB2		PB1	PA		HA		NP		NA			NS1	
	149	627	187	548	550	157*	127	277	340	9	435		171	
NJ 06	T	K	R	L	M	K	G	P	D	A	R		D	
NJ 01	P	E	K	M	L	E	E	H	N	T	K		N	
Avian	P(1099/1100) S(1/1100)	E(1018/1100) V(74/1100) K(6/1100) A(1/1100) G(1/1100)	R(1106/1108) K(2/1108)	M(1164/1168) I(3/1168) X(1/1168)	L(1119/1168) I(49/1168)	K(2017/2032) N(3/2032) Q(3/2032) R(2/2032) X(2/2032) E(2/2032) T(2/2032) D(1/2032)	E(1178/1182) D(2/1182) K(1/1182) G(1/1182)	P(1181/1182) H(1/1182)	D(1178/1182) G(1/1182) E(1/1182) H(1/1182) N(1/1182)	A(1210/1431) T(192/1431) V(15/1431) S(7/1431) P(2/1431) I(2/1431) E(1/1431) F(1/1431) Y(1/1431)	R(1392/1431) K(26/1431) I(7/1431) S(3/1431) G(3/1431)		D(1339/1663) N(93/1663) T(64/1663) E(54/1663) Y(39/1663) G(11/1663) A(1/1663) H(1/1663) V(1/1663)	
Swine	P(28/28)	E(39/39)	R(28/30) K(2/30)	M(31/31)	L(31/31)	K(43/44) N(1/44)	E(31/32) D(1/32)	P(35/35)	D(35/35)	A(29/32) T(3/32)	R(41/43) K(2/43)		D(40/42) N(2/42)	
Human	P(14/14)	E(13/13)	R(14/14)	M(14/14)	L(14/14)	K(21/21)	E(14/14)	P(13/13)	D(13/13)	A(18/18)	R(20/20)		D(11/14) E(3/14)	

**H3 numbering.*

are involved in the severe influenza in humans and animal models caused by avian viruses, such as H5N1 and H7N9, or highly pathogenic human viruses, such as the 1918 H1N1 virus (Gambotto et al., 2008; Franco-Paredes et al., 2009; Uyeki, 2009; Zheng et al., 2013; Feng et al., 2014; Li et al., 2014; Hrincius et al., 2015). Our results showed that the NJ06 infection could result in severe edema and alveolar damage, and elevated inflammatory cell infiltration in mouse lungs, whereas no difference in lung water content was observed between the NJ01-infected group and the control group, and only mild and limited alveolitis was observed in the NJ01-infected lungs. In addition, the NJ06 infection resulted in significantly higher numbers of inflammatory cells in BAL than NJ01 or mock infection. Furthermore, the percentages of neutrophils in BAL cells in NJ06-infected mice were significantly higher than those in NJ01-infected mice. Neutrophils are primary mediator/effector cells involved in producing acute lung injury (Headley et al., 1997; Ayala et al., 2002), and the elevated levels of neutrophils have also been found in the BAL samples of mice infected with highly pathogenic H5N1 viruses or the novel H7N9 viruses (Xu et al., 2009, 2013; Feng et al., 2015). Therefore, enhanced pulmonary neutrophil invasion may be associated with the severity of NJ06 virus infection in mice.

It is generally accepted that dysregulation of cytokine response is associated with the high virulence of AIVs in mammals (Us, 2008). As expected, the NJ06 virus caused intense expression of proinflammatory cytokine genes, such as TNF-α, CXCL10, and IL-17a. TNF-α is a key factor modulating neutrophil activity, and a high level of this cytokine has been linked to the hyperresponsiveness of neutrophils (Grommes and Soehnlein, 2011). CXCL10 is a potent chemoattractant for activated Th1 lymphocytes and natural killer cells and plays a role in the temporal development of innate and adaptive immunity in concert with type I and II IFNs (Neville et al., 1997). The high levels of TNF-α and CXCL10 have been linked to the persistent severe viral disease in patients with severe acute respiratory syndrome (Tobinick, 2004; de Jong et al., 2006). IL-17a acts as a pro-inflammatory cytokine that induces the expansion and accumulation of neutrophils of the innate immune system (Perrone et al., 2008; Crowe et al., 2009) and plays a critical role in mediating the acute lung injury caused by 2009 pandemic H1N1 influenza infection (Li et al., 2012). Therefore, based on the established role of these cytokines in viral disease, our results suggest that these pro-inflammatory cytokines may have pathological importance in NJ06 infection and are partially responsible for disease pathogenesis.

Although the NJ06 virus showed higher virulence and induced more severe lung injury in mice compared with the NJ01virus, the two viruses differed only by 12 amino acids distributed throughout seven genes. Except for PB2-E627K, none of these amino acid differences had been recognized to be related to increased virulence or replication efficiency. The amino acid at position 627 of PB2 is a well known determinant of host range, and the substitution E627K has been shown to be crucial for the adaptation and increased virulence of avian influenza viruses in mammals (Subbarao et al., 1993; Hatta et al., 2001; Fornek et al.,

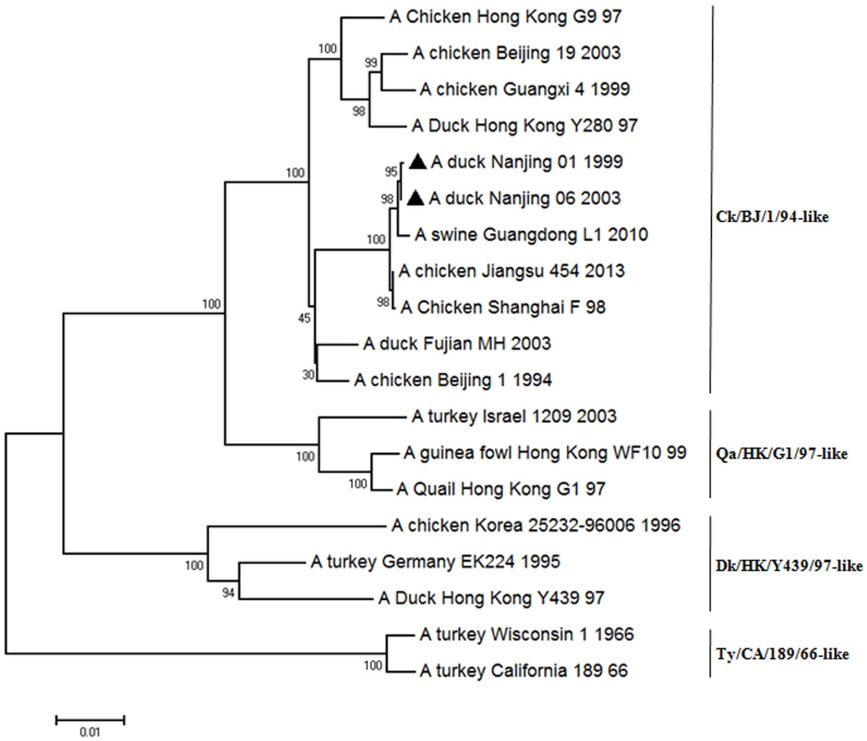


FIGURE 6 | Phylogenetic tree of HA genes of representative H9N2 influenza A viruses. The tree was generated by the distance-based neighbor-joining method using software MEGA 6.06. The reliability of the tree was assessed by bootstrap analysis with 1000 replicates. The viruses tested in this study are marked with black triangle.

2009; Li et al., 2009). However, H5N1 isolates with PB2 627E are also lethal to human (Mehle and Doudna, 2009) and mouse, whereas H9N2 isolates with PB2 627K are not lethal for mouse (Wang et al., 2012; Liu et al., 2014), indicating that the PB2 627K is not a sole determinant factor for mammalian adaptation by avian influenza viruses or its contribution to virulence need to interact with residue at other positions or genes (Wang et al., 2012; Liu et al., 2015). In addition, all other ten differences, except for NA-K435R, were located in the recognized functional regions that are involved in interaction of viral proteins or between the virus and host factors. The HA-E157K residues locate to the antigenic site I of the H9 HA corresponding to site B in H3 HA (Both et al., 1983; Kaverin et al., 2004), and the NA-T9A residues located in the amino-terminal transmembrane domain (Shtyrya et al., 2009). The NS-N171D residues resides in the host cleavage and polyadenylation factor (CPSF30) binding domain (Nemeroff et al., 1998), which has been found to be associated with the inhibition of 3′-end processing of cellular pre-mRNAs, including IFN-β pre-mRNA (Noah et al., 2003). Seven different residues (PB2-P149T, PB2-E627K, PB1-K187R, PA-M548L, PA-L550M, NP-E127G, NP-H277P, and NP-N340D) in the RNP complexes were all located in the known functional regions, including the NP binding domain of the PB2 protein, the nuclear localization sequence (NLS) of the PB1 protein, the PB1 binding domain of the PA protein, and the PB2 binding regions of the NP protein (Liu et al., 2009; Ng et al., 2009; Boivin et al., 2010; Ping et al., 2011). Therefore, besides PB2 E627K, other residue differences may also be associated with the high virulence of the NJ06 virus.

In summary, our study showed that a natural H9N2 isolate is highly pathogenic for mice, which might suggest the potential threat of H9N2 AIVs for other mammals, including humans, and also highlight the necessity for continued evaluation of the viral pathogenicity for mammals in the surveillance for LPAIVs, especially H9N2 viruses. In addition, comparison of the predicted amino acid sequences of NJ06 and NJ01 viruses showed that twelve residues differences in specific functional regions of viral genome resulted in the highly pathogenic phenotypes of the NJ06 virus, including rapid growth *in vivo* and *in vitro*, severe pulmonary lesions, excessive inflammatory cellular infiltration and cytokine response in lungs, and death in mice. However, it is not yet known which residues differences, or combinations of differences, are responsible for the high virulence for mice. We are currently attempting to determine the role of individual mutations in the viruses' pathogenicity in mice using a reverse genetics approach.

AUTHOR CONTRIBUTIONS

YiL and QL conceived and designed this research; QL, YuL, JY, and XH performed this experiments; DZ and KH analyzed

the data; KB contributed reagents/analytic tools; QL and YuL wrote the paper. All authors read and approved the paper.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (grant no. 31502100), and the Jiangsu Agricultural Science and Technology Innovation Fund [CX (15)1058].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01737/full#supplementary-material

REFERENCES

Alexander, D. J. (2000). A review of avian influenza in different bird species. *Vet. Microbiol.* 74, 3–13. doi: 10.1016/S0378-1135(00)00160-7

Alexander, D. J. (2007). An overview of the epidemiology of avian influenza. *Vaccine* 25, 5637–5644. doi: 10.1016/j.vaccine.2006.10.051

Ayala, A., Chung, C. S., Lomas, J. L., Song, G. Y., Doughty, L. A., Gregory, S. H., et al. (2002). Shock-induced neutrophil mediated priming for acute lung injury in mice: divergent effects of TLR-4 and TLR-4/FasL deficiency. *Am. J. Pathol.* 161, 2283–2294. doi: 10.1016/S0002-9440(10)64504-X

Belser, J. A., and Tumpey, T. M. (2013). H5N1 pathogenesis studies in mammalian models. *Virus Res.* 178, 168–185. doi: 10.1016/j.virusres.2013.02.003

Bi, J., Deng, G., Dong, J., Kong, F., Li, X., Xu, Q., et al. (2010). Phylogenetic and molecular characterization of H9N2 influenza isolates from chickens in Northern China from 2007-2009. *PLoS ONE* 5:e13063. doi: 10.1371/journal.pone.0013063

Boivin, S., Cusack, S., Ruigrok, R. W., and Hart, D. J. (2010). Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. *J. Biol. Chem.* 285, 28411–28417. doi: 10.1074/jbc.R110.117531R110.117531

Both, G. W., Sleigh, M. J., Cox, N. J., and Kendal, A. P. (1983). Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. *J. Virol.* 48, 52–60.

Butt, A. M., Siddique, S., Idrees, M., and Tong, Y. (2010). Avian influenza A (H9N2): computational molecular analysis and phylogenetic characterization of viral surface proteins isolated between 1997 and 2009 from the human population. *Virol. J.* 7:319. doi: 10.1186/1743-422X-7-319

Butt, K. M., Smith, G. J., Chen, H., Zhang, L. J., Leung, Y. H., Xu, K. M., et al. (2005). Human infection with an avian H9N2 influenza a virus in Hong Kong in 2003. *J. Clin. Microbiol.* 43, 5760–5767. doi: 10.1128/JCM.43.11.5760-5767.2005

Cauldwell, A. V., Long, J. S., Moncorge, O., and Barclay, W. S. (2014). Viral determinants of influenza a virus host range. *J. Gen. Virol.* 95(Pt 6), 1193–1210. doi: 10.1099/vir.0.062836-0

Chen, H., Deng, G., Li, Z., Tian, G., Li, Y., Jiao, P., et al. (2004). The evolution of H5N1 influenza viruses in ducks in southern China. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10452–10457. doi: 10.1073/pnas.0403212101

Choi, Y. K., Ozaki, H., Webby, R. J., Webster, R. G., Peiris, J. S., Poon, L., et al. (2004). Continuing evolution of H9N2 influenza viruses in Southeastern China. *J. Virol.* 78, 8609–8614. doi: 10.1128/JVI.78.16.8609-8614.2004

Crowe, C. R., Chen, K., Pociask, D. A., Alcorn, J. F., Krivich, C., Enelow, R. I., et al. (2009). Critical role of IL-17RA in immunopathology of influenza infection. *J. Immunol.* 183, 5301–5310. doi: 10.4049/jimmunol.0900995

de Jong, M. D., Simmons, C. P., Thanh, T. T., Hien, V. M., Smith, G. J., Chau, T. N., et al. (2006). Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nat. Med.* 12, 1203–1207. doi: 10.1038/nm1477

Densmore, J. C., Jeziorczak, P. M., Clough, A. V., Pritchard, KA Jr, Cummens, B., Medhora, M., et al. (2013). Rattus model utilizing selective pulmonary ischemia induces bronchiolitis obliterans organizing pneumonia. *Shock* 39, 271–277. doi: 10.1097/SHK.0b013e318281a58c

Feng, L., Wu, J. T., Liu, X., Yang, P., Tsang, T. K., Jiang, H., et al. (2014). Clinical severity of human infections with avian influenza A(H7N9) virus, China, 2013/14. *Euro Surveill.* 19:20984. doi: 10.2807/1560-7917.ES2014.19.49.20984

Feng, Y., Hu, L., Lu, S., Chen, Q., Zheng, Y., Zeng, D., et al. (2015). Molecular pathology analyses of two fatal human infections of avian influenza A(H7N9) virus. *J. Clin. Pathol.* 68, 57–63. doi: 10.1136/jclinpath-2014-202441

Fornek, J. L., Gillim-Ross, L., Santos, C., Carter, V., Ward, J. M., Cheng, L. I., et al. (2009). A single-amino-acid substitution in a polymerase protein of an H5N1 influenza virus is associated with systemic infection and impaired T-cell activation in mice. *J. Virol.* 83, 11102–11115. doi: 10.1128/JVI.00994-09

Franco-Paredes, C., Hernandez-Ramos, I., Del Rio, C., Alexander, K. T., Tapia-Conyer, R., and Santos-Preciado, J. I. (2009). H1N1 influenza pandemics: comparing the events of 2009 in Mexico with those of 1976 and 1918-1919. *Arch. Med. Res.* 40, 669–672. doi: 10.1016/j.arcmed.2009.10.004

Fusaro, A., Monne, I., Salviato, A., Valastro, V., Schivo, A., Amarin, N. M., et al. (2011). Phylogeography and evolutionary history of reassortant H9N2 viruses with potential human health implications. *J. Virol.* 85, 8413–8421. doi: 10.1128/JVI.00219-11

Gabriel, G., Dauber, B., Wolff, T., Planz, O., Klenk, H. D., and Stech, J. (2005). The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18590–18595. doi: 10.1073/pnas.0507415102

Gambotto, A., Barratt-Boyes, S. M., de Jong, M. D., Neumann, G., and Kawaoka, Y. (2008). Human infection with highly pathogenic H5N1 influenza virus. *Lancet* 371, 1464–1475. doi: 10.1016/S0140-6736(08)60627-3

Grommes, J., and Soehnlein, O. (2011). Contribution of neutrophils to acute lung injury. *Mol Med* 17, 293–307. doi: 10.2119/molmed.2010.00138

Guo, Y., Li, J., and Cheng, X. (1999). [Discovery of men infected by avian influenza a (H9N2) virus]. *Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi* 13, 105–108.

Guo, Y. J., Krauss, S., Senne, D. A., Mo, I. P., Lo, K. S., Xiong, X. P., et al. (2000). Characterization of the pathogenicity of members of the newly established H9N2 influenza virus lineages in Asia. *Virology* 267, 279–288. doi: 10.1006/viro.1999.0115

Hatta, M., Gao, P., Halfmann, P., and Kawaoka, Y. (2001). Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* 293, 1840–1842. doi: 10.1126/science.1062882

Headley, A. S., Tolley, E., and Meduri, G. U. (1997). Infections and the inflammatory response in acute respiratory distress syndrome. *Chest* 111, 1306–1321. doi: 10.1378/chest.111.5.1306

Hoffmann, E., Stech, J., Guan, Y., Webster, R. G., and Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza a viruses. *Arch. Virol.* 146, 2275–2289. doi: 10.1007/s007050170002

Homme, P. J., and Easterday, B. C. (1970). Avian influenza virus infections. I. characteristics of influenza a-turkey-Wisconsin-1966 virus. *Avian Dis.* 14, 66–74. doi: 10.2307/1588557

Hrincius, E. R., Liedmann, S., Finkelstein, D., Vogel, P., Gansebom, S., Ehrhardt, C., et al. (2015). Nonstructural protein 1 (NS1)-mediated inhibition of c-Abl results in acute lung injury and priming for bacterial co-infections: insights into 1918 H1N1 pandemic? *J. Infect. Dis.* 211, 1418–1428. doi: 10.1093/infdis/jiu609

Hu, J., Hu, Z., Song, Q., Gu, M., Liu, X., Wang, X., et al. (2013). The PA-gene-mediated lethal dissemination and excessive innate immune response contribute to the high virulence of H5N1 avian influenza virus in mice. *J. Virol.* 87, 2660–2672. doi: 10.1128/JVI.02891-12

Jia, N., Wang, S. X., Liu, Y. X., Zhang, P. H., Zuo, S. Q., Lin, Z., et al. (2008). Increased sensitivity for detecting avian influenza-specific antibodies by a modified hemagglutination inhibition assay using horse erythrocytes. *J. Virol. Methods* 153, 43–48. doi: 10.1016/j.jviromet.2008.06.016

Katz, J. M., Lu, X., Tumpey, T. M., Smith, C. B., Shaw, M. W., and Subbarao, K. (2000). Molecular correlates of influenza A H5N1 virus pathogenesis in mice. *J. Virol.* 74, 10807–10810. doi: 10.1128/JVI.74.22.10807-10810.2000

Kaverin, N. V., Rudneva, I. A., Ilyushina, N. A., Lipatov, A. S., Krauss, S., and Webster, R. G. (2004). Structural differences among hemagglutinins of influenza a virus subtypes are reflected in their antigenic architecture: analysis of H9 escape mutants. *J. Virol.* 78, 240–249. doi: 10.1128/JVI.78.1.240-249.2004

Kimble, J. B., Sorrell, E., Shao, H., Martin, P. L., and Perez, D. R. (2011). Compatibility of H9N2 avian influenza surface genes and 2009 pandemic H1N1 internal genes for transmission in the ferret model. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12084–12088. doi: 10.1073/pnas.1108058108

Lang, J. D., Figueroa, M., Sanders, K. D., Aslan, M., Liu, Y., Chumley, P., et al. (2005). Hypercapnia via reduced rate and tidal volume contributes to lipopolysaccharide-induced lung injury. *Am. J. Respir. Crit. Care Med.* 171, 147–157. doi: 10.1164/rccm.200302-305OC

Li, C., Yang, P., Sun, Y., Li, T., Wang, C., Wang, Z., et al. (2012). IL-17 response mediates acute lung injury induced by the 2009 pandemic influenza A (H1N1) virus. *Cell Res.* 22, 528–538. doi: 10.1038/cr.2011.165

Li, C., Yu, K., Tian, G., Yu, D., Liu, L., Jing, B., et al. (2005). Evolution of H9N2 influenza viruses from domestic poultry in Mainland China. *Virology* 340, 70–83. doi: 10.1016/j.virol.2005.06.025

Li, C., Zhang, A. J., To, K. K., Lee, A. C., Zhu, H., Wu, H. W., et al. (2014). Avian influenza A H7N9 virus induces severe pneumonia in mice without prior adaptation and responds to a combination of zanamivir and COX-2 inhibitor. *PLoS ONE* 9:e107966. doi: 10.1371/journal.pone.0107966

Li, J., Ishaq, M., Prudence, M., Xi, X., Hu, T., Liu, Q., et al. (2009). Single mutation at the amino acid position 627 of PB2 that leads to increased virulence of an H5N1 avian influenza virus during adaptation in mice can be compensated by multiple mutations at other sites of PB2. *Virus Res.* 144, 123–129. doi: 10.1016/j.virusres.2009.04.008

Li, Z., Jiang, Y., Jiao, P., Wang, A., Zhao, F., Tian, G., et al. (2006). The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. *J. Virol.* 80, 11115–11123. doi: 10.1128/JVI.00993-06

Lin, Y. P., Shaw, M., Gregory, V., Cameron, K., Lim, W., Klimov, A., et al. (2000). Avian-to-human transmission of H9N2 subtype influenza a viruses: relationship between H9N2 and H5N1 human isolates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9654–9658. doi: 10.1073/pnas.160270697

Liu, Q., Chen, H., Huang, J., Chen, Y., Gu, M., Wang, X., et al. (2014). A nonpathogenic duck-origin H9N2 influenza a virus adapts to high pathogenicity in mice. *Arch. Virol.* 159, 2243–2252. doi: 10.1007/s00705-014-2062-y

Liu, Q., Huang, J., Chen, Y., Chen, H., Li, Q., He, L., et al. (2015). Virulence determinants in the PB2 gene of a mouse-adapted H9N2 virus. *J. Virol.* 89, 877–882. doi: 10.1128/JVI.01775-14

Liu, Y., Lou, Z., Bartlam, M., and Rao, Z. (2009). Structure-function studies of the influenza virus RNA polymerase PA subunit. *Sci. China C Life Sci.* 52, 450–458. doi: 10.1007/s11427-009-0060-1

Long, J. S., Benfield, C. T., and Barclay, W. S. (2015). One-way trip: influenza virus' adaptation to gallinaceous poultry may limit its pandemic potential. *Bioessays* 37, 204–212. doi: 10.1002/bies.201400133

Mehle, A., and Doudna, J. A. (2009). Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21312–21316. doi: 10.1073/pnas.0911915106

Mok, C. K., Yen, H. L., Yu, M. Y., Yuen, K. M., Sia, S. F., Chan, M. C., et al. (2011). Amino acid residues 253 and 591 of the PB2 protein of avian influenza virus A H9N2 contribute to mammalian pathogenesis. *J. Virol.* 85, 9641–9645. doi: 10.1128/JVI.00702-11

Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W., and Krug, R. M. (1998). Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol. Cell* 1, 991–1000. doi: 10.1016/S1097-2765(00)80099-4

Neville, L. F., Mathiak, G., and Bagasra, O. (1997). The immunobiology of interferon-gamma inducible protein 10 kD (IP-10): a novel, pleiotropic member of the C-X-C chemokine superfamily. *Cytokine Growth Factor Rev.* 8, 207–219. doi: 10.1016/S1359-6101(97)00015-4

Ng, A. K., Wang, J. H., and Shaw, P. C. (2009). Structure and sequence analysis of influenza A virus nucleoprotein. *Sci. China C Life Sci.* 52, 439–449. doi: 10.1007/s11427-009-0064-x

Nick, J. A., Young, S. K., Brown, K. K., Avdi, N. J., Arndt, P. G., Suratt, B. T., et al. (2000). Role of p38 mitogen-activated protein kinase in a murine model of pulmonary inflammation. *J. Immunol.* 164, 2151–2159. doi: 10.4049/jimmunol.164.4.2151

Noah, D. L., Twu, K. Y., and Krug, R. M. (2003). Cellular antiviral responses against influenza a virus are countered at the posttranscriptional level by the viral NS1A protein via its binding to a cellular protein required for the 3' end processing of cellular pre-mRNAS. *Virology* 307, 386–395. doi: 10.1016/S0042-6822(02)00127-7

Oh, D. Y., and Hurt, A. C. (2016). Using the ferret as an animal model for investigating influenza antiviral effectiveness. *Front. Microbiol.* 7:80. doi: 10.3389/fmicb.2016.00080

Panwen, B. S., Zhang, Z. F., Fan, X. H., Liu, J. Y., Lai, Z. P., Luo, J. L., et al. (2012). [Serological survey on antibody to H9 subtypes of avian influenza a virus in medical college students in Guangxi of China and Hue of Vietnam]. *Zhonghua Liu Xing Bing Xue Za Zhi* 33, 444–445.

Peiris, J. S., Guan, Y., Markwell, D., Ghose, P., Webster, R. G., and Shortridge, K. F. (2001). Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses in pigs in southeastern China: potential for genetic reassortment? *J. Virol.* 75, 9679–9686. doi: 10.1128/JVI.75.20.9679-9686.2001

Peiris, M., Yuen, K. Y., Leung, C. W., Chan, K. H., Ip, P. L., Lai, R. W., et al. (1999). Human infection with influenza H9N2. *Lancet* 354, 916–917. doi: 10.1016/S0140-6736(99)03311-5

Perk, S., Panshin, A., Shihmanter, E., Gissin, I., Pokamunski, S., Pirak, M., et al. (2006). Ecology and molecular epidemiology of H9N2 avian influenza viruses isolated in Israel during 2000-2004 epizootic. *Dev. Biol. (Basel)* 124, 201–209.

Perrone, L. A., Plowden, J. K., Garcia-Sastre, A., Katz, J. M., and Tumpey, T. M. (2008). H5N1 and 1918 pandemic influenza virus infection results in early and excessive infiltration of macrophages and neutrophils in the lungs of mice. *PLoS Pathog.* 4:e1000115. doi: 10.1371/journal.ppat.1000115

Ping, J., Keleta, L., Forbes, N. E., Dankar, S., Stecho, W., Tyler, S., et al. (2011). Genomic and protein structural maps of adaptive evolution of human influenza a virus to increased virulence in the mouse. *PLoS ONE* 6:e21740. doi: 10.1371/journal.pone.0021740

Salomon, R., Franks, J., Govorkova, E. A., Ilyushina, N. A., Yen, H. L., Hulse-Post, D. J., et al. (2006). The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. *J. Exp. Med.* 203, 689–697. doi: 10.1084/jem.20051938

Shtyrya, Y. A., Mochalova, L. V., and Bovin, N. V. (2009). Influenza virus neuraminidase: structure and function. *Acta Naturae* 1, 26–32.

Sirinonthanawech, N., Uiprasertkul, M., Suptawiwat, O., and Auewarakul, P. (2011). Viral load of the highly pathogenic avian influenza H5N1 virus in infected human tissues. *J. Med. Virol.* 83, 1418–1423. doi: 10.1002/jmv.22146

Song, M. S., Pascua, P. N., Lee, J. H., Baek, Y. H., Lee, O. J., Kim, C. J., et al. (2009). The polymerase acidic protein gene of influenza a virus contributes to pathogenicity in a mouse model. *J. Virol.* 83, 12325–12335. doi: 10.1128/JVI.01373-09

Subbarao, E. K., London, W., and Murphy, B. R. (1993). A single amino acid in the PB2 gene of influenza a virus is a determinant of host range. *J. Virol.* 67, 1761–1764.

Sun, H., Cui, P., Song, Y., Qi, Y., Li, X., Qi, W., et al. (2015). PB2 segment promotes high-pathogenicity of H5N1 avian influenza viruses in mice. *Front. Microbiol.* 6:73. doi: 10.3389/fmicb.2015.00073

Sun, X., Xu, X., Liu, Q., Liang, D., Li, C., He, Q., et al. (2013). Evidence of avian-like H9N2 influenza a virus among dogs in Guangxi. *China Infect. Genet. Evol.* 20, 471–475. doi: 10.1016/j.meegid.2013.10.012

Thangavel, R. R., and Bouvier, N. M. (2014). Animal models for influenza virus pathogenesis, transmission, and immunology. *J. Immunol. Methods* 410, 60–79. doi: 10.1016/j.jim.2014.03.023

Thulasi Raman, S. N., and Zhou, Y. (2016). Networks of host factors that interact with NS1 protein of influenza a virus. *Front. Microbiol.* 7:654. doi: 10.3389/fmicb.2016.00654

Tobinick, E. (2004). TNF-alpha inhibition for potential therapeutic modulation of SARS coronavirus infection. *Curr. Med. Res. Opin.* 20, 39–40. doi: 10.1185/030079903125002757

Us, D. (2008). [Cytokine storm in avian influenza]. *Mikrobiyol. Bul.* 42, 365–380.

Uyeki, T. M. (2009). Human infection with highly pathogenic avian influenza A (H5N1) virus: review of clinical issues. *Clin. Infect. Dis.* 49, 279–290. doi: 10.1086/600035

Wan, H., and Perez, D. R. (2007). Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. *J. Virol.* 81, 5181–5191. doi: 10.1128/JVI.02827-06

Wan, H., Sorrell, E. M., Song, H., Hossain, M. J., Ramirez-Nieto, G., Monne, I., et al. (2008). Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. *PLoS ONE* 3:e2923. doi: 10.1371/journal.pone.0002923

Wang, J., Sun, Y., Xu, Q., Tan, Y., Pu, J., Yang, H., et al. (2012). Mouse-adapted H9N2 influenza a virus PB2 protein M147L and E627K mutations are critical for high virulence. *PLoS ONE* 7:e40752. doi: 10.1371/journal.pone.0040752

Wang, M., Fu, C. X., and Zheng, B. J. (2009). Antibodies against H5 and H9 avian influenza among poultry workers in China. *N. Engl. J. Med.* 360, 2583–2584. doi: 10.1056/NEJMc0900358

Xu, L., Bao, L., Deng, W., Zhu, H., Chen, T., Lv, Q., et al. (2013). The mouse and ferret models for studying the novel avian-origin human influenza A (H7N9) virus. *Virol. J.* 10, 253. doi: 10.1186/1743-422X-10-253

Xu, T., Qiao, J., Zhao, L., He, G., Li, K., Wang, J., et al. (2009). Effect of dexamethasone on acute respiratory distress syndrome induced by the H5N1 virus in mice. *Eur. Respir. J.* 33, 852–860. doi: 10.1183/09031936.00130507

Yu, H., Zhou, Y. J., Li, G. X., Ma, J. H., Yan, L. P., Wang, B., et al. (2011). Genetic diversity of H9N2 influenza viruses from pigs in China: a potential threat to human health? *Vet. Microbiol.* 149, 254–261. doi: 10.1016/j.vetmic.2010.11.008

Zhang, Z., Hu, S., Li, Z., Wang, X., Liu, M., Guo, Z., et al. (2011). Multiple amino acid substitutions involved in enhanced pathogenicity of LPAI H9N2 in mice. *Infect. Genet. Evol.* 11, 1790–1797. doi: 10.1016/j.meegid.2011.07.025

Zheng, Y., Cao, Y., Lu, Y., Xi, X., Qian, Z., Lowrie, D., et al. (2013). Clinical comparison of one recovered case and one fatal case of human infection with H7N9 avian influenza in Shanghai Public Health Clinical Center in China. *Malays J. Med. Sci.* 20, 76–79.

Zhou, H., Yu, Z., Hu, Y., Tu, J., Zou, W., Peng, Y., et al. (2009). The special neuraminidase stalk-motif responsible for increased virulence and pathogenesis of H5N1 influenza a virus. *PLoS ONE* 4:e6277. doi: 10.1371/journal.pone.0006277

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Liu, Liu, Yang, Huang, Han, Zhao, Bi and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 12 January 2017
doi: 10.3389/fmicb.2016.02130

Influenza A Viruses Replicate Productively in Mouse Mastocytoma Cells (P815) and Trigger Pro-inflammatory Cytokine and Chemokine Production through TLR3 Signaling Pathway

Di Meng[1†], Caiyun Huo[1†], Ming Wang[1,2], Jin Xiao[1,2], Bo Liu[1], Tangting Wei[1], Hong Dong[3], Guozhong Zhang[1], Yanxin Hu[1] and Lunquan Sun[4]**

[1] Key Laboratory of Animal Epidemiology and Zoonosis of Ministry of Agriculture, College of Veterinary Medicine, China Agricultural University, Beijing, China, [2] Key Laboratory of Veterinary Bioproduction and Chemical Medicine of the Ministry of Agriculture, Zhongmu Institutes of China Animal Husbandry Industry Co., Ltd, Beijing, China, [3] Beijing Key Laboratory of Traditional Chinese Veterinary Medicine, Beijing University of Agriculture, Beijing, China, [4] Center for Molecular Medicine, Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Noemi Sevilla,
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Spain
Qinfang Liu,
Shanghai Veterinary Research Institute, China

***Correspondence:**
Yanxin Hu
huyx@cau.edu.cn
Lunquan Sun
lunquansun @csu.edu.cn

[†] These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 07 May 2016
Accepted: 16 December 2016
Published: 12 January 2017

Citation:
Meng D, Huo C, Wang M, Xiao J, Liu B, Wei T, Dong H, Zhang G, Hu Y and Sun L (2017) Influenza A Viruses Replicate Productively in Mouse Mastocytoma Cells (P815) and Trigger Pro-inflammatory Cytokine and Chemokine Production through TLR3 Signaling Pathway. Front. Microbiol. 7:2130. doi: 10.3389/fmicb.2016.02130

The influenza A viruses (IAVs) cause acute respiratory infection in both humans and animals. As a member of the initial lines of host defense system, the role of mast cells during IAV infection has been poorly understood. Here, we characterized for the first time that both avian-like (α-2, 3-linked) and human-like (α-2, 6- linked) sialic acid (SA) receptors were expressed by the mouse mastocytoma cell line (P815). The P815 cells did support the productive replication of H1N1 (A/WSN/33), H5N1 (A/chicken/Henan/1/04) and H7N2 (A/chicken/Hebei/2/02) *in vitro* while the *in vivo* infection of H5N1 in mast cells was confirmed by the specific staining of nasal mucosa and lung tissue from mice. All the three viruses triggered the infected P815 cells to produce pro-inflammatory cytokines and chemokines including IL-6, IFN-γ, TNF-α, CCL-2, CCL-5, and IP-10, but not the antiviral type I interferon. It was further confirmed that TLR3 pathway was involved in P815 cell response to IAV-infection. Our findings highlight the remarkable tropism and infectivity of IAV to P815 cells, indicating that mast cells may be unneglectable player in the development of IAV infection.

Keywords: influenza A viruses, mast cells, pro-inflammatory cytokines, chemokines, TLR3 pathway

INTRODUCTION

Influenza A virus (IAV) is one of the most common respiratory pathogen in humans and animals. Since the first outbreak in Hong Kong in 1997, the highly pathogenic avian influenza (HPAI) H5N1 virus has become a public health threat due to its potential to cause serious illness and death in humans (Uyeki, 2009). Virus-induced acute lung injury or its more severe form, acute respiratory distress syndrome (ARDS), is a major cause of mortality by pandemic influenza and HAPI H5N1 infections; however, the exact mechanism of this injury is not fully understood. Several studies suggest that the main contributing factor is an increased production of inflammatory

cytokines or "cytokine storm" (Thuy et al., 2011). This response may result from each individual cell producing more cytokines, or through chemokines-induced recruitment of a greater number of innate immune cells into the lung (Teijaro et al., 2011). Thus, the cellular sources involved in the resulting cytokine storm remain undetermined.

Mast cells are enriched near epithelial surfaces exposed to the external environment, and thus function as sentinels in the defense against host infection, they also play a role in initiating adaptive immune responses (Shelburne and Abraham, 2011). These processes are aided by the expression of a unique 'armamentarium' of receptor systems and mediators for responding to pathogen-associated signals (Marshall, 2004). Mast cells are crucial for optimal immune responses against bacterial, parasitic, and viral infections (Marshall, 2004). It was well demonstrated that mast cells played important roles in the pathogenesis of some viral infections, such as HIV-1, dengue virus, cytomegalovirus and bovine respiratory syncytial virus (Gibbons et al., 1990; King et al., 2000; Jolly et al., 2004; Sundstrom et al., 2004; Shirato and Taguchi, 2009). We previously demonstrated that mast cells were activated by H5N1 virus infection and escalated lung injury (Hu et al., 2012). However, it remains to be determined how this response is initiated, and whether IAV can infect and efficiently replicate in mast cells.

Influenza viruses bind to neuraminic acids (sialic acids, SA) on the surface of cells to initiate infection and replication (Knipe et al., 2001). The expression of SA linkages is cell type specific. For example, α-2,3-SA receptors are detected specifically in ciliated cells, while α-2,6-SA receptors are exclusively present in non-ciliated cells (Matrosovich et al., 2004). Historically, α-2,6-linked SA receptors, which are preferentially recognized by human influenza viruses, are detected exclusively in cells of upper respiratory tract of humans. However, both the α-2,6-linkages and α-2,3- linkages are present in the human lower respiratory tract, predominantly recognized by human and avian influenza viruses respectively (Raman et al., 2014). Thus, the type and distribution of SA is an important determinant of influenza virus tropism and pathogenesis (Suzuki et al., 2000); yet, little is known about SA receptor expression on mast cells.

During IAV infection, influenza viral dsRNA is sensed through several classes of pattern-recognition receptors (PRRs) including Toll-like receptors (TLRs) and retinoic acid-inducible gene-I-like receptors (RLRs) (Yu and Levine, 2011). Among the PPRs, toll-like receptor 3 (TLR3) and cytolytic RNA helicases retinoic acid-inducible gene I (RIG-I) are the most common transducers of viral dsRNA signals. Once stimulated with their respective agonist, TLR3 recruits adaptor molecule TIR-domain containing adaptor inducing interferon-β (TRIF), while RIG-I associated with mitochondrial antiviral signaling (MAVS) to initiate downstream signaling. Both these pathways activate the transcription factor nuclear factor (NF)-kB, leading to the production of inflammatory cytokines, chemokines, and activation of interferon regulatory factor (IRF) 3 and/or 7 to induce the key antiviral mediator type I IFNs (Takeuchi and Akira, 2009). Recently, Le Goffic et al. (2007) demonstrated the importance of TLR3 in the inflammatory cytokine response to IAV in lung epithelial cells *in vitro* (Le Goffic et al., 2007). IAV-TLR3 interactions are also critical for viral pathology *in vivo*. Previous studies showed that the association between RIG-I and viral ssRNA bearing an uncapped 5′-triphosphate end (Pichlmair et al., 2006) and this association resulted in the production of IFNs (Kato et al., 2006). Moreover, RIG-I played a key role in the expression of proinflammatory cytokines in mast cells infected by IAV. However, the role of TLR3 during IAV infection in mast cells remains unexplored.

Therefore, in the present study we sought to determine the presence and role of SA receptors on mouse mastocytoma cell line (P815). We demonstrated that P815 cells expressed both α-2, 3-, and α-2, 6- linked SA receptors to initiate IAV infection. In addition, P815 cells supported productive replication of IAVs *in vitro* while the *in vivo* infection of H5N1 in mast cells was confirmed by the specific staining of nasal mucosa and lung tissue from mice. Following IAV infection, P815 cells mediated substantial hyper-induction of pro-inflammatory cytokines and chemokines, and TLR3 signal pathways probably involved in the process. This provides insight for the development of novel strategies to combat influenza infection by targeting mast cells.

MATERIALS AND METHODS

Ethics Statement

All mouse experimental protocols complied with the guidelines of the Beijing Laboratory Animal Welfare and Ethics Committee, and were approved by the Beijing Association for Science and Technology (the approval ID is SYXK-2009-0423). All experiments with live H5/H7 subtype viruses were performed in a biosafety level 3 containment laboratory (the approval number is CNAS BL0017) approved by the Ministry of Agriculture of the People's Republic of China.

Viruses and Cells Culture

The avian influenza viruses H5N1 (A/Chicken/Henan/1/04) (Hu et al., 2012) and H7N2 (A/Chicken/Hebei/2/02) were isolated from infected chicken flocks, and propagated in the allantoic cavities of 10-day-old embryonated chicken eggs for 24–48 h at 37°C. The working stocks of human influenza virus H1N1 (A/WSN/33) were generated in MDCK cells. Virus titers were determined by standard plaque assays. The 50% lethal dose (LD_{50}) in mice was determined as previously described (Hu et al., 2012). The mouse mastocytoma cell line P815 and the Madin-Darby canine kidney cell line MDCK were cultured as previously described (Hu et al., 2012).

In vitro Viral Infection and LE-PolyI:C Treatment

Cell monolayers were formed in tissue culture plates by seeding 6-well (1×10^6 cells/well) or 12-well (5×10^5 cells/well) plates, washed with DMEM and infected with viruses at a multiplicity of infection (MOI) of 0.1 for 1 h at 37°C. After incubation, cells monolayers were washed and DMEM supplemented with 1%

bovine serum albumin was added to each well and incubated for the indicated times. Polyinosine-polycytidylic acid (polyI:C), a synthetic mimic of viral double-stranded RNA, was used as a positive control. Liposome-encapsulated PolyI:C (LE-PolyI:C) used in this study was prepared as described previously (Wong et al., 1999), diluted to a final concentration of 10 μg/ml and incubated with cells at 37°C for the indicated times.

In vivo Viral Challenge

Female BALB/c mice (8–10 weeks) were purchased from Vital River Laboratories (Beijing, China), and feed pathogen-free food and water in independent ventilated cages. Mice were first anesthetized with Zotile® (Virbac, Carros, France), and then infected intra-nasally with PBS-diluted H5N1 virus (5LD50) or PBS alone. The nose and lung tissues were then collected 6 days post-infection.

Immunofiuorescence Staining and Confocal Microscopy

Tissue samples were fixed in 4% neutral formalin, embedded in paraffin, and serially cut at a thickness of 5 μm. Cultured cells were fixed on a polylysine-coated slide with 4% formaldehyde, and blocked with 3% BSA. To visualize surface receptors, slides containing fixed tissues or cells were directly stained with fluorescein- *Sambucus nigra* bark lectin (SNA, specific to SAα2,6-Gal) or fluorescein- *Maackia amurensis* lectin I (MAA-I, specific to SAα2,3-Galβ(1-4)GlcNAc). To confirm the specificity of lectin binding, monolayers were washed and treated with 250 mU/ml of neuraminidase from *Clostridium perfringens* (New England BioLabs, Beijing, China) for 3 h prior to lectin staining. To detect tryptase expression or IAV nucleoprotein (NP) antigen, cells were permeabilized with 0.5% Triton X-100 before blocking, then tissue sections or cell slides were either incubated with a rabbit anti-mast cell tryptase monoclonal antibody (Abcam, [EPR8476], ab134932) for 2 h at room temperature, or a mouse anti-IAV NP monoclonal antibody (Abcam [AA5H], ab20343) at 4°C overnight. After washing three times with PBS-T, tissue sections were further incubated with a Texas red-conjugated goat anti-mouse or rabbit secondary antibody, and cell slides were incubated with a FITC-conjugated goat anti-mouse secondary antibody (Abcam) for 1 h at room temperature. To visualize the nuclei, all slides were stained with 3 μg/ml 4′,6′-Diamidine-2-phenylindole (DAPI) (Sigma–Aldrich) for 5 min at room temperature and then examined under a laser scanning confocal microscope (Leica TCS SP5 II, Leica Microsystems, Wetzlar, Germany).

Flow Cytometry Analysis

Cultured P815 cells (1×10^6) were pelleted, washed twice with DMEM, once with flow buffer (PBS with 2% FBS) and then re-suspended in 200 μl of fluorescein- SAA or fluorescein- MAA I at different dilutions. The cells were incubated for 1 h at 4°C, then washed and re-suspended in flow cytometry buffer for analysis on a BD FACSCalibur using Cell Questpro software (BD Biosciences, California).

Transmission Electronmicroscopy (TEM)

Cells were trypsinized and fixed using 2.5% (v/v) glutaraldehyde in PBS for 2 h at 4°C. Cells were then washed with PBS, post-fixed in 1% osmium tetroxide, and washed and dehydrated in series of ethanol solutions. The dehydrated pellets were embedded in epoxy resin, and 70-nm sections were cut. Then the sections were placed on copper sieves and stained with uranyl acetate and lead citrate. Images were obtained using a JEM-1230 TEM (JEOL, Japan Electronics Co., Ltd, Tokyo, Japan).

Real-Time Quantitative PCR

Total RNA was extracted from cells in Trizol reagents (Invitrogen, Carlsbad, CA, USA). DNase I-treated RNA (0.2 μg) was reverse transcribed into cDNA using random or universal primers for IAV (Uni 12) (Hoffmann et al., 2001) with an EasyScript First-Strand cDNA Synthesis Super Mix (TransGen Biotech, China) according to the manufacturer's instruction. Reactions were performed in triplicate using a Power SYBR® Green PCR Master Mix (Applied Biosystems, Warrington, UK) and the Applied Biosystems 7500 system. The mRNA expression levels of the genes were normalized to β-actin, compared with mock-infected cells, and quantified by the $2^{-\Delta\Delta CT}$ method. The sequences of the primers are listed in Supplementary Table S1. The amplifications were performed as follows: a 10 min hot start at 95°C, followed by 40 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 35 s, and extension at 72°C for 40 s.

Proteome Profiler Antibody Array Assay

Cell-free supernatants were acquired by centrifugation, then the levels of cytokines and chemokines were analyzed by the Mouse Cytokine Array Panel A (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's instructions, which could provide parallel determination of the relative levels of 40 kinds of selected mouse cytokines and chemokines.

Cytokine and Chemokine Quantification

The concentration of IL-6, IFN-γ, RANTES, IP-10, IFN-α, IFN-β, and TNF-α in the supernatant of cell cultures was determined using ELISA kits (eBioscience, San Diego, CA, USA) according to the manufacturer's instructions.

Co-immunoprecipitation

Cells were either mock-treated, LE-PolyI:C-treated, or infected with IAV viruses at a MOI of 0.1. At 6 h post-infection, the cells were washed with cold PBS and lysed for 15 min on ice with RIPA lysis buffer containing 50 mM Tris-HCl (pH7.4), 150 mM NaCl, 1% NP40, 0.25% sodium deoxycholate and a protease inhibitor cocktail (Beyotime Institute of Biotechnology, Beijing, China). Lysed cells were pelleted and the supernatants were incubated with the indicated antibodies (anti-TLR3 antibody or an isotype IgG antibody from Abcam) for 2 h at 4°C with gentle shaking. The samples were then added to a preprocessed EZview™ Red Protein A/G Affinity Gel (Sigma) and incubated for 6 h. After washing three times with lysis buffer, the beads were boiled in SDS loading buffer, and then analyzed by immunoblot with the indicated

antibodies (anti-TLR3 antibody or anti-TRIF antibody from Abcam).

Western Blot

Cell lysates were prepared using lysis buffer as described above, and protein concentrations were determined using a BCA protein assay kit (Beyotime Institute of Biotechnology). Equal amounts of protein were separated by SDS-PAGE and transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore, Beijing, China). The membranes were blocked using 5% non-fat dry milk (BD Biosciences) at room temperature for 2 h, and then incubated overnight at 4°C with antibodies (anti-TLR3 antibody and anti-TRIF antibody from Abcam; Influenza A NS1 antibody from Santa Cruz Biotechnology, Dallas, Texas, USA). After three 10 min washes in Tris-buffered saline containing 0.05% Tween (TBST), the membranes were incubated for 1 h at room temperature with the appropriate horseradish peroxidase-conjugated secondary antibody. Protein bands were visualized using the Western Lightning Plus-ECL (Perkin Elmer, MA, USA). β- Actin was used as a loading control.

Inhibition of TLR3 Activation Using Specific Inhibitions

P815 cell monolayers were pre-incubated with TLR3/dsRNA complex inhibitor (Calbiochem, Darmstadt, Germany) at a concentration of 25 and 50 μM (diluted with DMSO) for 12 h. They were then were infected with virus at a MOI of 0.1, exposed to LE-PolyI:C, or mock treated as described above. The same concentration of inhibitors was immediately added after 1 h of viral incubation. Samples were collected 24 and 36 h after infection.

Statistical Analysis

Statistical analysis was performed by one-way ANOVA using the SPSS software suite (version 12.0), and a P-value of <0.05 was considered statistically significant. Results were expressed as mean ± standard deviation (SD) of at least three independent experiments.

RESULTS

P815 Cells Express Both α-2,3- Linked and α-2,6- Linked SA Receptors

Influenza viral HA proteins initiate infection by interacting with sialic acid residues coating the surface of host cells. In general, human influenza viruses and *S. nigra* bark lectin (SNA) preferentially bind α-2,6- linked SA receptors, while avian influenza viruses and *M. amurensis* lectin (MAA) predominantly bind to α-2,3- linked SA receptors (Springer et al., 1969; Suzuki et al., 2000; Ibricevic et al., 2006; Shinya et al., 2006). Here we used the mouse mastocytoma cell line, P815, as a mast cell model. To determine the susceptibility of P815 cells to influenza viruses, we first analyzed the distribution of surface SA receptors by lectin staining as described in the materials and methods. Both α-2,3- and α-2,6- linked SA receptors were expressed on the surface of P815 mouse mastocytoma cells, with the intensities of SNA being visually stronger than one of the two isoforms of MAA (**Figure 1A**). We treated P815 cells with a broad-spectrum neuraminidase to cleave sialic moieties abolished lectin binding, and confirmed the specificity of SNA and MAA staining (**Figure 1A**, insets).

To quantitatively analyze the sialic acid residues, P815 cells were stained with different concentrations of FITC-conjugated lectins and assayed by flow cytometry. As shown in **Figure 1B**, at an lectin concentration of 5 μg/ml, 95.54% of P815 cells were detected positive expression of α-2,3-linked SA receptors (FITC-MAA) and 98.57% of α-2,6-linked receptors (FITC-SNA). At higher concentrations (10 and 20 μg/ml), almost all cells positively expressed both SA receptors (>99%). The mean fluorescence intensity (MFI) depended on the lectin concentration, and the MFI of SNA was significantly higher (>fivefold) than that of MAA at all the concentrations. These data suggested that the expression of α-2,6- linked SA receptors was more abundant on P815 mastocytoma cells than α-2,3- linked SA receptors.

Taken together, these data suggested that both α-2,3- and α-2,6- linked SA receptors were expressed on the surface of P815 mouse mastocytoma cells.

P815 Cells Support the Replications of IAVs

We previously demonstrated that H5N1 infection could activate mast cells (Hu et al., 2012). To determine if IAV could enter and replicate productively in mastocytoma cells, we examined the replication kinetics of human and avian IAVs in P815 cells. As shown in Supplementary Figure S1, all three subtypes of IAVs productively replicated in P815 cells as measured by hemagglutination assay (left), plaque formation (middle), and viral RNA expression (right). In P815 cells, the replications H1N1 and H5N1 viruses were more efficient than H7N2 virus. These data indicated that IAVs replicated well in P815 cells with some degree of tropism selectivity.

To corroborate this finding, P815 cells were co-stained with α-2,3- or α-2,6- linked SA receptors and viral NP (**Figure 2A**). The wide distribution of α-2,6- linked SA receptors was consistent the effective replication of H1N1. In addition, a large number of NP-positive cells were observed when infected with H5N1, but were less abundant in H7N2 infected cells, which was consistent with the viral replication profiles (Supplementary Figure S1). To further validate the permissiveness of P815 for IAVs replication, we used transmission electron microscopy. As shown in **Figure 2B**, budding virions were present on the surface of cells infected with each of the three virus subtypes. Moreover, many viral particles were found to be associated with the surface of the cells infected with H1N1 and H5N1, but were less obvious on H7N2 infected cells. Together, the above data suggested that IAVs could bind and enter into

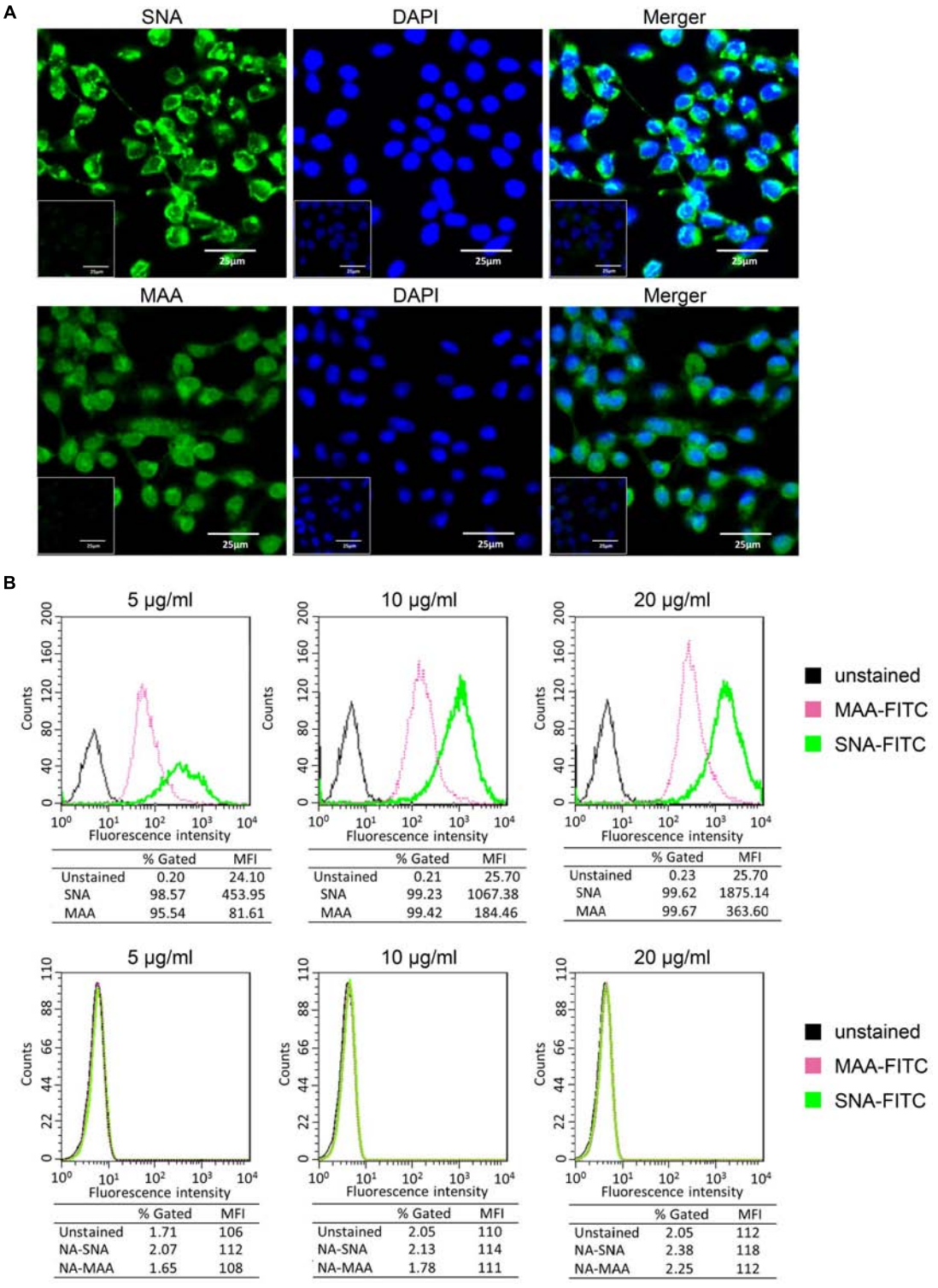


FIGURE 1 | P815 cells express α-2, 3- and α-2,6- linked sialic acid (SA) receptors. (A) The mouse mastocytoma cell line P815 was placed on polylysine-coated slides and stained with FITC-conjugated SNA or MAA-I (green), and DAPI (blue) for nuclei. Image inserts depict cells pre-treated with neuraminidase to abolished sialic acid residue staining. **(B)** Trypsinized P815 cells were incubated with FITC-conjugated SNA or MAA-I (concentrations from left to right are 5, 10, and 20 μg/ml) and analyzed using flow cytometry to determine relative percentages of cells expressing α-2,3-SA (MAA, pink) or α-2,6-SA (SNA, green), compared to unstained cells (black). "NA-MAA" and "NA-SNA" indicated that P815 cells pre-treated with neuraminidase to abolished sialic acid residue staining. Results shown are representative of three independent experiments.

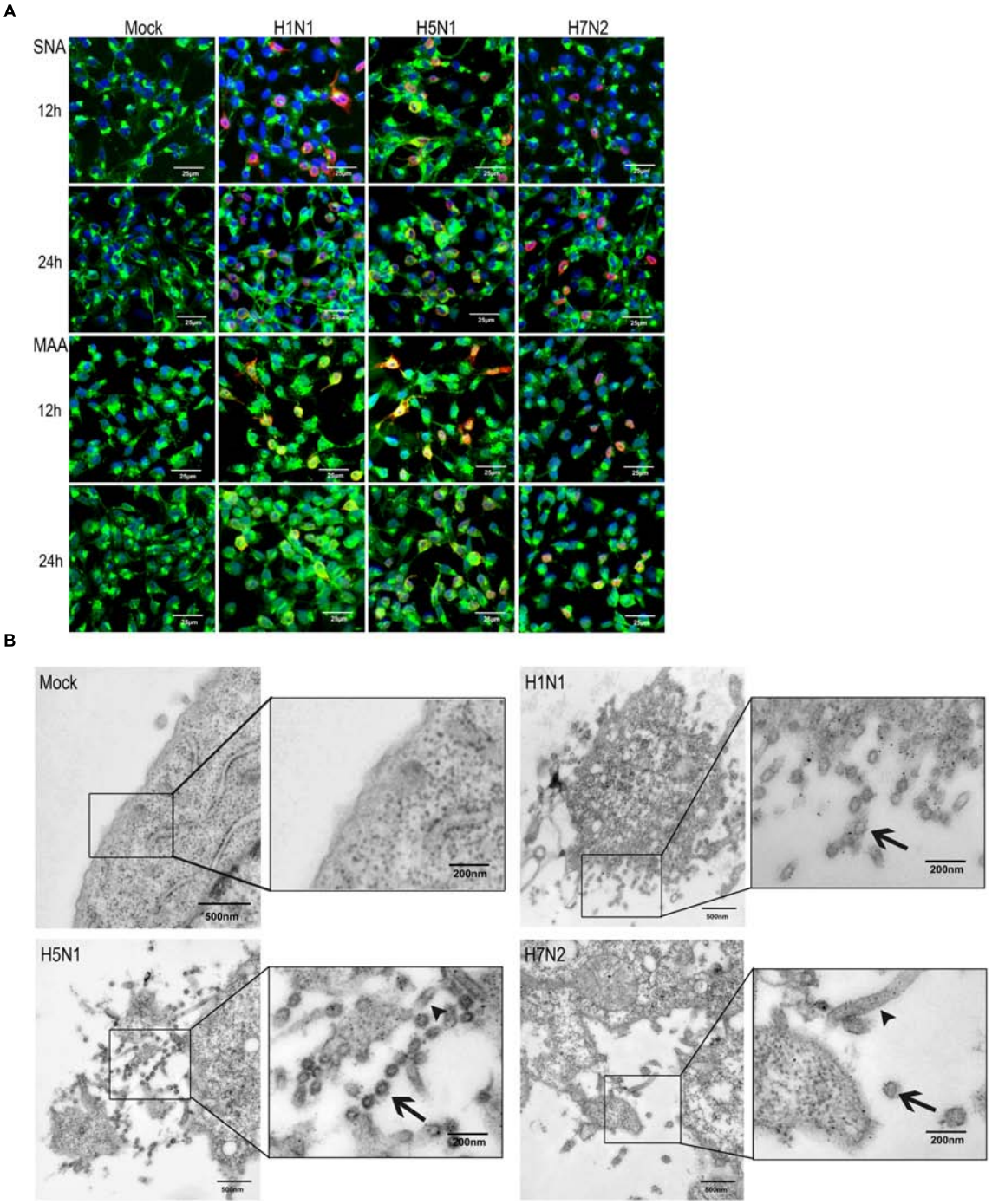


FIGURE 2 | Influenza viruses can infect P815 cells *in vitro*. P815 cells were mock-treated or infected with three subtypes of influenza viruses at a MOI of 0.1. **(A)** Immunofluorescence detection of viral nucleoprotein (NP) antigen and SA receptors in P815 cells. At the indicated times post-infection, cells were fixed and stained for α-2,6- or α-2,3- sialic acids (green) and influenza NP (red). Results shown are representative of three separate experiments. **(B)** Transmission electron microscopy of influenza viruses released from the cell surface. Higher magnifications are in the boxes on the right. Arrows denote the virus particles.

P815 mastocytoma cells, where the efficient replication was supported.

To determine if IAVs can infect mast cells *in vivo*, we used a mouse H5N1 virus infection model. Mice were either infected with H5N1 or PBS and the nasal mucosa and lung tissues were resected and analyzed by immunofluorescence. Cells double positive for the virus-specific antigen NP and the mast cell specific protein tryptase were present in H5N1 infected mice but

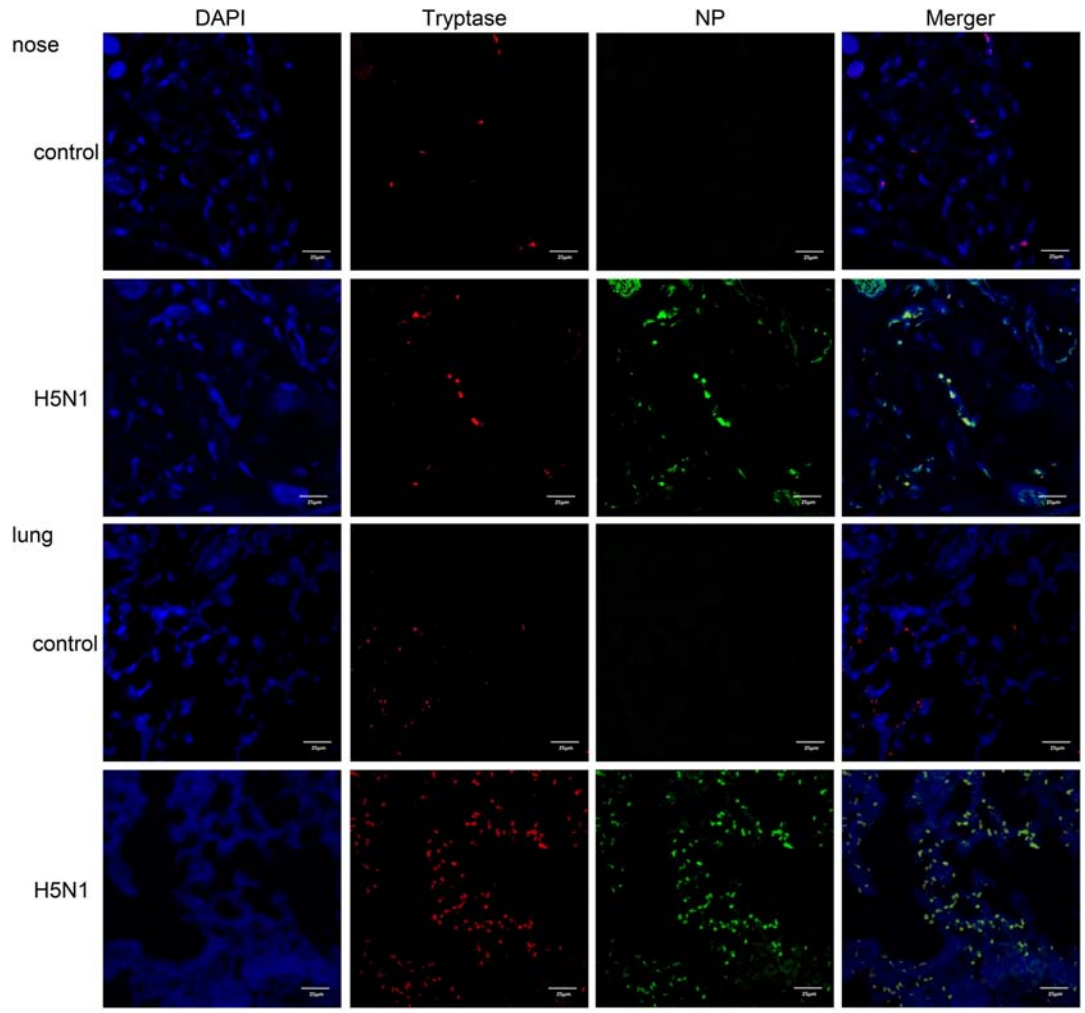


FIGURE 3 | H5N1 influenza virus can infect mouse mast cells *in vivo*. Representative nasal mucosa and lung sections from control or H5N1 virus-infected mice. Sections were analyzed by immunofluorescence staining (Blue = nuclei, red = tryptase, and green = influenza NP). Results shown are representative of three independent repeats.

not in PBS treated mice (**Figure 3**). These data suggested that IAVs probably could actively infect mast cells *in vivo*.

IAVs Induce Cytokine and Chemokine Production in P815 Cells

Our previous data suggested that H5N1-activated mast cells could intensify lung injury by releasing the pro-inflammatory mediators including histamine, tryptase, and IFN-γ (Hu et al., 2012). To more specifically examine the cytokines and chemokines released by P815 cells upon avian and human IAVs infection, we performed an antibody array analysis. As shown in **Figure 4A**, the production of sICAM-1, IL-6, IL-13, IP-10, M-CSF, CCL-2, CCL-12, and TNF-α was augmented in the supernatants of P815 cells infected with all three viruses subtypes. The release of G-CSF, GM-CSF, CCL-1, IFN-γ, IL-1α, IL-4, and CXCL-1 was only moderately increased. To further confirm these findings, we used ELISA to conduct kinetic profiles of the selected cytokines and chemokines that were potentially involved in responses to influenza infection. The three virus subtypes, and LE-PolyI:C, induced significantly higher levels of IL-6 and IFN-γ compared to mock treated cells (**Figure 4B**). The production of CCL-2 did not occur until 12 h post-infection, but then increased from 24 to 48 h post-infection (**Figure 4B**), while the secretion of IP-10 and TNF-α increased gradually peaking at 36 h post-infection. In contrast, all three IAV subtypes induced a relatively low expression of CCL-5 for 24 h post-infection, but this increased at later time points. The expressions of antiviral cytokines IFN-α and IFN-β did not change from 2 to 48 h post-infection (**Figure 4B**).

To further analyze the expression kinetics of various cytokines and chemokines released by P815 cells infected with each of the three virus subtypes and LE-PolyI:C, we performed quantitative RT-PCR. The mRNA expression profiles of IL-6, IFN-γ, TNF-α, CCL-2, CCL-5, and IP-10 were similar to the data generated from

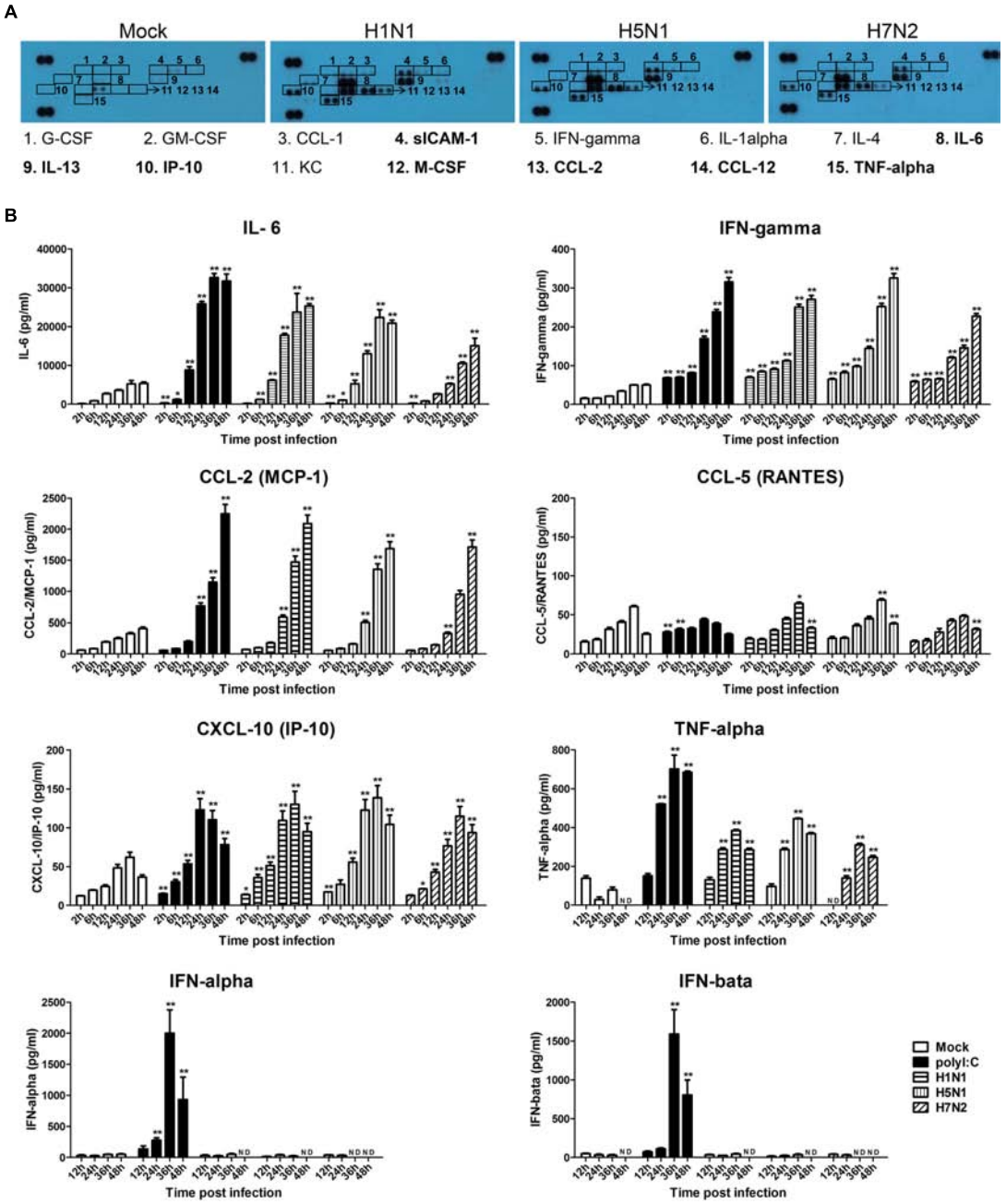


FIGURE 4 | Influenza virus infection causes the release of pro-inflammatory cytokines and chemokines. (A) P815 cells were infected with H1N1, H5N1, and H7N2 at a MOI of 0.1, or mock treated for 24 h. Cell-free supernatants were then analyzed for cytokine and chemokine content using a cytokine array panel. Numbered boxes denote up-regulated expression, with the most dramatic increases annotated in bold. Results shown are representative of two separate experiments. **(B)** P815 cells were infected with H1N1, H5N1, and H7N2 at a MOI of 0.1, exposed to LE-PolyI:C, or mock treated. At the designated time points, cell supernatants were harvested and the expression of IL-6, IFN-γ, CCL-2, CCL-5, IP-10 TNF-α, IFN-α, and IFN-β was analyzed by ELISA. Graphs shown are mean ± SD of three independent replicates. Asterisks indicate statistically significant increases compared to mock treatment (*$P < 0.05$, **$P < 0.01$). ND, not detectable.

ELISA (Supplementary Table S2). However, while the expression levels of a large number of pro-inflammatory cytokines and chemokines were up-regulated, the mRNA levels of the antiviral genes IFN-α, IFN-β and the anti-inflammatory cytokine IL-10 were unchanged during viral infection (Supplementary Table S2). Taken together, these data suggested that following IAV infection P815 cells released a range of pro-inflammatory cytokines and chemokines.

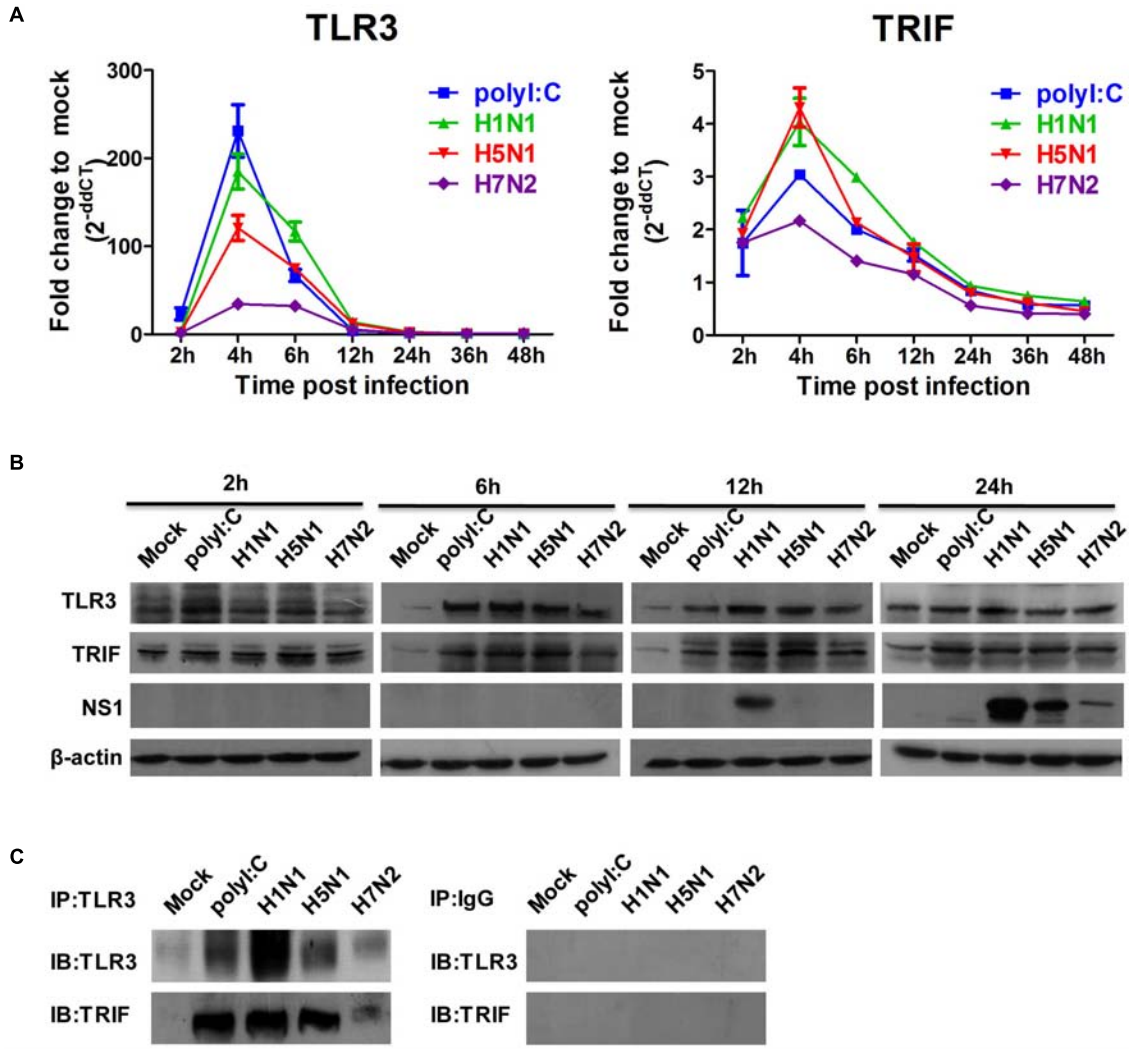


FIGURE 5 | Influenza A viruses up-regulates the expressions of TLR3 and TRIF mRNA and proteins. P815 cells were treated or infected as described in **Figure 4B**. **(A)** Total RNA was isolated at the designed times and examined with quantitative real-time PCR. The expression of TLR3and TRIF is shown. The data are presented as the relative fold change over mock treatment, and are pooled from three independent experiments. **(B)** P815 cells were collected at 2, 6, 12, and 24 h post-infection or treatment, and then analyzed by immunoblotting for expression of TLR3, TRIF, viral NS1 protein and β-actin. Data are representative of three separate experiments. **(C)** Endogenous interactions of TLR3 with TRIF. Whole cell extracts at 6 h (TLR3-TRIF) post-infection were immunoprecipitated with the indicated antibodies or isotype IgG controls and analyzed by western blot analysis.

TLR3 Plays a Key Role in the Expression of Proinflammatory Cytokines in P815 Cells Infected by IAV

Given that IAVs could infect P815 cells and promote the release of inflammatory cytokines and chemokines, we next examined expression level of viral RNA sensor TLR3 involved in the transduction of inflammatory molecule signals. The mRNA expression level of TLR3 in P815 cells infected with IAVs peaked 4 h post-infections, and returned to baseline levels by 12 h post-infection (**Figure 5A**). In addition, the mRNA expression profiles of the adaptor molecules TRIF increased in a manner consistent with TLR3 levels; however, the fold increase compared to mock infected cells was much lower. Consistent with mRNA levels, the protein expression of TRL3 and TRIF was strongly augmented in IAV infected P815 cells (**Figure 5B**). To confirm if TLR3 were indeed activated, co-immunoprecipitation experiments were used to test the endogenous interactions of TLR3 with TRIF. Co-precipitation of TLR3/TRIF was evident in P815 cells infected with all three IAVs (**Figure 5C**). These data indicated that the viral RNA sensors TLR3 were expressed following IAV infection.

To further investigate the role of TLR3 in P815 cells during IAV infection, we utilized a novel TLR3/dsRNA complex inhibitor to disrupt the interaction between IAV and TLR3 in P815 cells. The effects of the inhibitors on the release of pro-inflammatory cytokines and chemokines, including IL-6, IFN-γ,

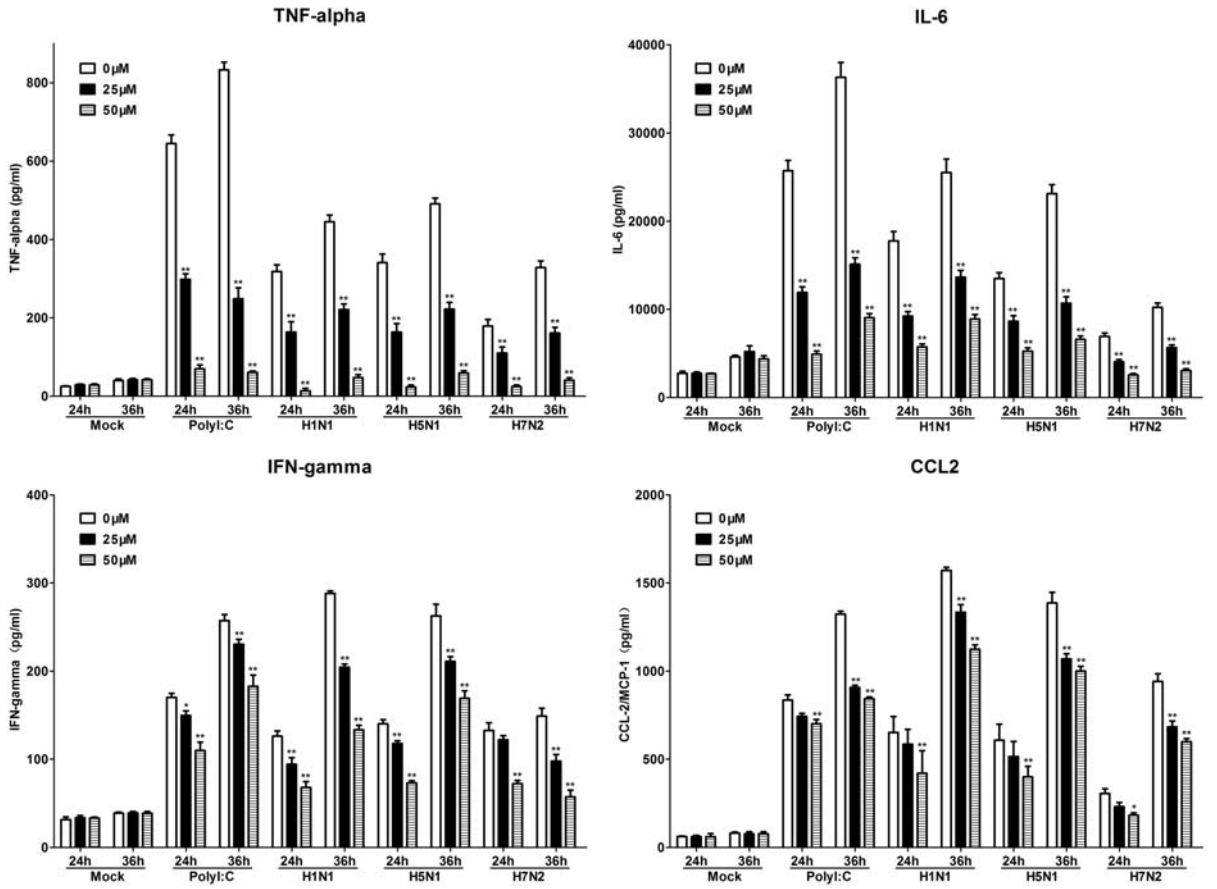


FIGURE 6 | Toll-like receptors/dsRNA complex inhibitor can reduce the releasing of proinflammatory cytokines in P815 cells infected by IAV. P815 cells treatment with TLR3/dsRNA complex inhibitor were infected with H1N1, H5N1, and H7N2 at a MOI of 0.1, exposed to LE-PolyI:C, or mock treated. At the designated time points, cell supernatants were harvested and the expression of TNF-α, IL-6, IFN-γ, and CCL-2 was analyzed by ELISA. Graphs shown are mean ± SD of three independent replicates. Asterisks indicate statistically significant increases compared to mock treatment (*$P < 0.05$, **$P < 0.01$).

TNF-α, and CCL-2, were examined. The expression of these pro-inflammatory cytokines and chemokines was significantly decreased in TLR3/dsRNA complex inhibitor treatment group 24 and 36 h after infection (**Figure 6**). In addition, treatment of IAV-infected P815 cells with TLR3/dsRNA complex inhibitor dramatically decreased viral titers 24 and 36 h after infection (**Figure 7**). Taken together, these data suggested that following IAV infection P815 cells actively participated in promoting inflammation by releasing a range of pro-inflammatory cytokines and chemokines possibly through TLR3 signaling pathways.

DISCUSSION

Elucidating the mechanisms of immune defense against IAV is critical for developing therapeutic strategies to prevent influenza infection. The role of endothelial cells, macrophages, and dendritic cells in preventing infection in the respiratory tract has been described (Bender et al., 1998; Julkunen et al., 2000). However, mast cells, an important cell type in the first lines of host immunity and defense, have been largely overlooked until recently, data from our lab and others have demonstrated a possible involvement of mast cells in IAV infection (Hu et al., 2012; Graham et al., 2013; Marcet et al., 2013). Our previous study showed that mast cells actively participate in the first-line immunological responses to IAV infection. Mast cells could aggravate pathological injury of the H5N1 virus infected tissues in mice by directly inducing apoptosis or inflammatory cytokines and mediators (Hu et al., 2012). However, the receptor repertoire facilitating IAV infection and the cellular response in mast cells are still largely unknown. The present work demonstrated that the mouse mastocytoma cell line (P815) expressed both α-2,3- and α-2,6- linked SA receptors to initiate IAV entry and could serve as a comfortable environment for virus replication and release progeny viruses.

In order to infect cells, IAV must first attach to SA receptors on the plasma membrane. We demonstrated that both α-2,3- and α-2,6- linked SA receptors were expressed on the surface of P815 mouse mastocytoma cells. To our knowledge, this is the first time SA receptors have been reported on the surface of cells considered as mast cells. In addition, our data demonstrated that P815 cells supported the productive replication of IAV. All

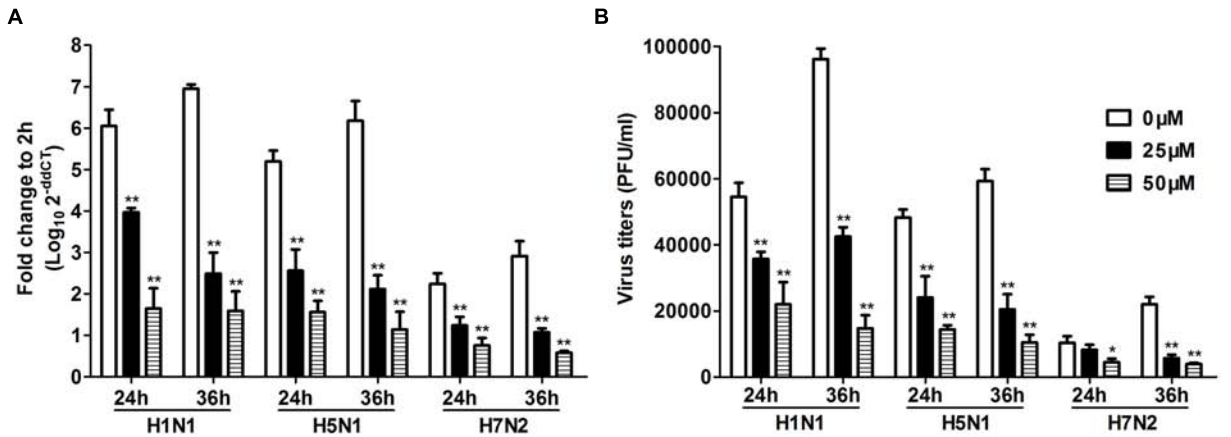


FIGURE 7 | Toll-like receptors/dsRNA complex inhibitor can decrease the viral loads of IAV in P815 cells. P815 cells treatment with TLR3/dsRNA complex inhibitor were infected with H1N1, H5N1, and H7N2 at a MOI of 0.1, exposed to LE-Polyl:C, or mock treated. Cells were homogenized in Trizol and relative viral NS gene quantification was determined by real time PCR **(A)**. Culture supernatants were collected at the indicated times post-infection, and virus titers in the supernatants were determined by plaque assay **(B)**. Results shown are pooled from three independent repeats. Asterisks indicate statistically significant increases compared to mock treatment (*$P < 0.05$, **$P < 0.01$).

three subtypes of the viruses (H1N1, H5N1, and H7N2) infected and replicated in P815 cells *in vitro*, with images of transmission electron microscope as the most powerful evidences. Though the mouse mastocytoma cell line, P815, as a cell model, has been widely used in mast cell-based research (Ohtsu et al., 1996; Lunderius et al., 2000; Zhang et al., 2009, 2010), the validity of data from this model should be verified by *in vivo* study or with primary cells. In our present study, the infection of mast cells with H5N1 in mice (*in vivo*) provided further evidence. Contradictory with the results in P815 mouse mastocytoma cells, Marcet's study showed that H1N1 virus were limited replication in human mast cells (Marcet et al., 2013). Similarly with our results in the present study, several groups showed that dengue virus could infect and replicate within mast cells (St John et al., 2011). Whether mast cells can be infected and serviced as a potent reservoir for persistent Human Immunodeficiency Virus was still a debate until now (Sundstrom et al., 2007). Thus, our findings provided new insights into the role of mast cells for the pathogenesis of influenza.

Activated mast cells can selectively release many pro-inflammatory cytokines and chemokines, which can vary greatly depending on the stimulus and experimental conditions (Abraham and St John, 2010). Here, we found that all the three subtypes of IAVs infected P815 mouse mastocytoma cells produced several cytokines, including a significant increase in the expression of IL-6, IFN-γ, and TNF-α, which could serve as objective markers of the host inflammatory responses and lung injury (Hierholzer et al., 1998; Perrone et al., 2010). Additionally, the chemokines CCL-2 and IP-10 were also increased, which was consistent with observations from both clinical settings and animal models (Zhou et al., 2006). Our results were consistent with a recent report that bone marrow-cultured mast cells infected by the influenza strain A/WSN/33 were found to release several mediators (Graham et al., 2013). Our findings showed that the three subtypes of IAVs induced similar cytokine and chemokine kinetics, and the magnitude of responses induced by H7N2 was much lower than H1N1 and H5N1, which exactly related to the viral replication competence but not the origin of the virus. We also found that the kinetic expression patterns of gene transcription were different for various cytokines and chemokines in P815 cells, some of them were likely directly induced by IAV infection as the time courses of their appearance were early, while others might result from autocrine or paracrine feedback as the late appearance. Interestingly, even though type I interferons are key antiviral cytokines produced by IAV-infected epithelial cells and monocytes/macrophage (Hofmann et al., 1997; Ronni et al., 1997), neither IFN-α nor IFN-β was detected in P815 cells. We speculate that mast cells are involved in the immune and inflammatory responses to AIV infection, but may not directly participate in the antiviral mechanisms.

Mast cells used TLR3, RIG-I and MDA5 to sense viral RNA following infection (Kulka et al., 2004; Oldstone and Rosen, 2014; Becker et al., 2015). In a Newcastle disease virus infection model, mast cells produced cytokines and chemokines in a TLR3-dependent manner (To et al., 2001). In addition, though the degranulation and the generation of eicosanoid in mast cells augmented the vascular leakage and played an important role in dengue virus infection (St John et al., 2013; Syenina et al., 2015), the roles of pattern recognition receptors against the viral infection were also indispensable. Moreover, mast cells infected with dengue virus and vesicular stomatitis virus were shown to activate RIG-I and MDA5, and produce cytokines and chemokines (Jacobs and Langland, 1996; St John et al., 2011; Teijaro et al., 2011; Brown et al., 2012). Importantly, Graham et al. (2013) found the inflammatory response of mast cells during IAV infection occurred in a RIG-I-dependent mechanism. However, the mechanism by which mast cells sense IAV infection is not well understood. Here, for the first time we demonstrated that

P815 mouse mastocytoma cells sensed influenza viruses mainly via TLR3. In the present study, we found that blocking TLR3 with TLR3/dsRNA complex inhibitor in the P815 cells resulted in decreased pro-inflammatory factors and viral titers. This data suggested that activated TLR3 pathways were responsible for the production of key pro-inflammatory cytokines and chemokines in IAV infected mast cells. A recent study suggested that TLR3 could act as viral sensor to mediate viral transactivation via upregulation of transcription factors such as c-Jun, which was known to regulate the viral promoter activity (Bhargavan et al., 2016). We recently also showed that IAV infection up-regulated c-jun expression and activation (Xie et al., 2014). This could partially explain that blocking TLR3 resulted in decreased IAV growth.

Collectively, our data suggest that mast cells not only participate in the IAV-induced immune response and inflammation, but also actively serve as reservoirs for IAV replication. Combined with we previously found that mast cells escalated lung injury could be reduced dramatically by treatment with ketotifen, which is a mast cell degranulation inhibitor (Hu et al., 2012; Han et al., 2016). Thus, considering the critical role of mast cells in IAV infection, this study provides insight for the development of novel strategies to combat influenza infection by targeting mast cells.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: DM, CH, YH, MW, and LS. Performed the experiments: DM, CH, TW, BL, JX, and HD. Analyzed the data: DM, CH, YH, and LS. Contributed reagents/materials/analysis tools: YH, MW, LS, GZ, and HD. Wrote the paper: DM, CH, YH, and LS. All authors reviewed the manuscript.

FUNDING

Research reported in this publication was supported by the National Twelve-five Technological Supported Plan of China (Grant no: 2015BAD12B01), the National Natural Science Foundation of China (Grant no. 31272531) and the Open Project Program of Beijing Key Laboratory of Traditional Chinese Veterinary Medicine at Beijing University of Agriculture (No. kf2016031).

ACKNOWLEDGMENT

The authors would like to thank all of the staff of the Key Laboratory of Animal Epidemiology and Zoonosis of Ministry of Agriculture, China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.02130/full#supplementary-material

REFERENCES

Abraham, S. N., and St John, A. L. (2010). Mast cell-orchestrated immunity to pathogens. *Nat. Rev. Immunol.* 10, 440–452. doi: 10.1038/Nri2782

Becker, M., Lemmermann, N. A. W., Ebert, S., Baars, P., Renzaho, A., Podlech, J., et al. (2015). Mast cells as rapid innate sensors of cytomegalovirus by TLR3/TRIF signaling-dependent and -independent mechanisms. *Cell. Mol. Immunol.* 12, 192–201. doi: 10.1038/Cmi.2014.73

Bender, A., Albert, M., Reddy, A., Feldman, M., Sauter, B., Kaplan, G., et al. (1998). The distinctive features of influenza virus infection of dendritic cells. *Immunobiology* 198, 552–567. doi: 10.1016/S0171-2985(98)80078-8

Bhargavan, B., Woollard, S. M., and Kanmogne, G. D. (2016). Toll-like receptor-3 mediates HIV-1 transactivation via NFkappaB and JNK pathways and histone acetylation, but prolonged activation suppresses Tat and HIV-1 replication. *Cell. Signal.* 28, 7–22. doi: 10.1016/j.cellsig.2015.11.005

Brown, M. G., McAlpine, S. M., Huang, Y. Y., Haidl, I. D., Al-Afif, A., Marshall, J. S., et al. (2012). RNA sensors enable human mast cell anti-viral chemokine production and IFN-mediated protection in response to antibody-enhanced dengue virus infection. *PLoS ONE* 7:e34055. doi: 10.1371/journal.pone.0034055

Gibbons, A. E., Price, P., Robertson, T. A., Papadimitriou, J. M., and Shellam, G. R. (1990). Replication of murine cytomegalovirus in mast cells. *Arch. Virol.* 115, 299–307. doi: 10.1007/BF01310538

Graham, A. C., Hilmer, K. M., Zickovich, J. M., and Obar, J. J. (2013). Inflammatory response of mast cells during influenza A virus infection is mediated by active infection and RIG-I signaling. *J. Immunol.* 190, 4676–4684. doi: 10.4049/jimmunol.1202096

Han, D., Wei, T., Zhang, S., Wang, M., Tian, H., Cheng, J., et al. (2016). The therapeutic effects of sodium cromoglycate against influenza A virus H5N1 in mice. *Influenza Other Respir. Viruses* 10, 57–66. doi: 10.1111/irv.12334

Hierholzer, C., Kalff, J. C., Omert, L., Tsukada, K., Loeffert, J. E., Watkins, S. C., et al. (1998). Interleukin-6 production in hemorrhagic shock is accompanied by neutrophil recruitment and lung injury. *Am. J. Physiol.* 275, L611–L621.

Hoffmann, E., Stech, J., Guan, Y., Webster, R. G., and Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* 146, 2275–2289. doi: 10.1007/s007050170002

Hofmann, P., Sprenger, H., Kaufmann, A., Bender, A., Hasse, C., Nain, M., et al. (1997). Susceptibility of mononuclear phagocytes to influenza A virus infection and possible role in the antiviral response. *J. Leukoc. Biol.* 61, 408–414.

Hu, Y., Jin, Y., Han, D., Zhang, G., Cao, S., Xie, J., et al. (2012). Mast cell-induced lung injury in mice infected with H5N1 influenza virus. *J. Virol.* 86, 3347–3356. doi: 10.1128/JVI.06053-11

Ibricevic, A., Pekosz, A., Walter, M. J., Newby, C., Battaile, J. T., Brown, E. G., et al. (2006). Influenza virus receptor specificity and cell tropism in mouse and human airway epithelial cells. *J. Virol.* 80, 7469–7480. doi: 10.1128/JVI.02677-05

Jacobs, B. L., and Langland, J. O. (1996). When two strands are better than one: the mediators and modulators of the cellular responses to double-stranded RNA. *Virology* 219, 339–349. doi: 10.1006/viro.1996.0259

Jolly, S., Detilleux, J., and Desmecht, D. (2004). Extensive mast cell degranulation in bovine respiratory syncytial virus-associated paroxystic respiratory distress syndrome. *Vet. Immunol. Immunopathol.* 97, 125–136. doi: 10.1016/j.vetimm.2003.08.014

Julkunen, I., Melen, K., Nyqvist, M., Pirhonen, J., Sareneva, T., and Matikainen, S. (2000). Inflammatory responses in influenza A virus infection. *Vaccine* 19, S32–S37. doi: 10.1016/S0264-410X(00)00275-9

Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., et al. (2006). Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101–105. doi: 10.1038/nature04734

King, C. A., Marshall, J. S., Alshurafa, H., and Anderson, R. (2000). Release of vasoactive cytokines by antibody-enhanced dengue virus infection of a human mast cell/basophil line. *J. Virol.* 74, 7146–7150. doi: 10.1128/JVI.74.15.7146-7150.2000

Knipe, D. M., Howley, P. M., and Griffin, D. E. (2001). *Fundamental Virology*. Philadelphia, PA: Lippincott Williams & Wilkins.

Kulka, M., Alexopoulou, L., Flavell, R. A., and Metcalfe, D. D. (2004). Activation of mast cells by double-stranded RNA: evidence for activation through toll-like receptor 3. *J. Allergy Clin. Immunol.* 114, 174–182. doi: 10.1016/j.jaci.2004.03.049

Le Goffic, R., Pothlichet, J., Vitour, D., Fujita, T., Meurs, E., Chignard, M., et al. (2007). Cutting Edge: influenza A virus activates TLR3-dependent inflammatory and RIG-I-dependent antiviral responses in human lung epithelial cells. *J. Immunol.* 178, 3368–3372. doi: 10.4049/jimmunol.178.6.3368

Lunderius, C., Xiang, Z., Nilsson, G., and Hellman, L. (2000). Murine mast cell lines as indicators of early events in mast cell and basophil development. *Eur. J. Immunol.* 30, 3396–3402. doi: 10.1002/1521-4141(2000012)30:12<3396::AID-IMMU3396>3.0.CO;2-O

Marcet, C. W., St Laurent, C. D., Moon, T. C., Singh, N., and Befus, A. D. (2013). Limited replication of influenza A virus in human mast cells. *Immunol. Res.* 56, 32–43. doi: 10.1007/s12026-012-8377-4

Marshall, J. S. (2004). Mast-cell responses to pathogens. *Nat. Rev. Immunol.* 4, 787–799. doi: 10.1038/nri1460

Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A., and Klenk, H. D. (2004). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4620–4624. doi: 10.1073/pnas.0308001101

Ohtsu, H., Kuramasu, A., Suzuki, S., Igarashi, K., Ohuchi, Y., Sato, M., et al. (1996). Histidine decarboxylase expression in mouse mast cell line P815 is induced by mouse peritoneal cavity incubation. *J. Biol. Chem.* 271, 28439–28444. doi: 10.1074/jbc.271.45.28439

Oldstone, M. B. A., and Rosen, H. (2014). Cytokine storm plays a direct role in the morbidity and mortality from influenza virus infection and is chemically treatable with a single sphingosine-1-phosphate agonist molecule. *Curr. Top. Microbiol. Immunol.* 378, 129–147. doi: 10.1007/978-3-319-05879-5_6

Perrone, L. A., Szretter, K. J., Katz, J. M., Mizgerd, J. P., and Tumpey, T. M. (2010). Mice lacking both TNF and IL-1 receptors exhibit reduced lung inflammation and delay in onset of death following infection with a highly virulent H5N1 virus. *J. Infect. Dis.* 202, 1161–1170. doi: 10.1086/656365

Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., et al. (2006). RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* 314, 997–1001. doi: 10.1126/science.1132998

Raman, R., Tharakaraman, K., Shriver, Z., Jayaraman, A., Sasisekharan, V., and Sasisekharan, R. (2014). Glycan receptor specificity as a useful tool for characterization and surveillance of influenza A virus. *Trends Microbiol.* 22, 632–641. doi: 10.1016/j.tim.2014.07.002

Ronni, T., Matikainen, S., Sareneva, T., Melen, K., Pirhonen, J., Keskinen, P., et al. (1997). Regulation of IFN-alpha/beta, MxA, 2',5'-oligoadenylate synthetase, and HLA gene expression in influenza A-infected human lung epithelial cells. *J. Immunol.* 158, 2363–2374.

Shelburne, C. P., and Abraham, S. N. (2011). The mast cell in innate and adaptive immunity. *Adv. Exp. Med. Biol.* 716, 162–185. doi: 10.1007/978-1-4419-9533-9_10

Shinya, K., Ebina, M., Yamada, S., Ono, M., Kasai, N., and Kawaoka, Y. (2006). Influenza virus receptors in the human airway. *Nature* 440, 435–436. doi: 10.1038/440435a

Shirato, K., and Taguchi, F. (2009). Mast cell degranulation is induced by A549 airway epithelial cell infected with respiratory syncytial virus. *Virology* 386, 88–93. doi: 10.1016/j.virol.2009.01.011

Springer, G. F., Schwick, H. G., and Fletcher, M. A. (1969). Relationship of influenza virus inhibitory activity of glycoproteins to their molecular size and sialic acid content. *Proc. Natl. Acad. Sci. U.S.A.* 64, 634–641. doi: 10.1073/pnas.64.2.634

St John, A. L., Rathore, A. P., Raghavan, B., Ng, M. L., and Abraham, S. N. (2013). Contributions of mast cells and vasoactive products, leukotrienes and chymase, to dengue virus-induced vascular leakage. *Elife* 2:e00481. doi: 10.7554/eLife.00481

St John, A. L., Rathore, A. P., Yap, H., Ng, M. L., Metcalfe, D. D., Vasudevan, S. G., et al. (2011). Immune surveillance by mast cells during dengue infection promotes natural killer (NK) and NKT-cell recruitment and viral clearance. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9190–9195. doi: 10.1073/pnas.1105079108

Sundstrom, J. B., Ellis, J. E., Hair, G. A., Kirshenbaum, A. S., Metcalfe, D. D., Yi, H., et al. (2007). Human tissue mast cells are an inducible reservoir of persistent HIV infection. *Blood* 109, 5293–5300. doi: 10.1182/blood-2006-11-058438

Sundstrom, J. B., Little, D. M., Villinger, F., Ellis, J. E., and Ansari, A. A. (2004). Signaling through Toll-like receptors triggers HIV-1 replication in latently infected mast cells. *J. Immunol.* 172, 4391–4401. doi: 10.4049/jimmunol.172.7.4391

Suzuki, Y., Ito, T., Suzuki, T., Holland, R. E., Chambers, T. M., Kiso, M., et al. (2000). Sialic acid species as a determinant of the host range of influenza A viruses. *J. Virol.* 74, 11825–11831. doi: 10.1128/Jvi.74.24.11825-11831.2000

Syenina, A., Jagaraj, C. J., Aman, S. A., Sridharan, A., and St John, A. L. (2015). Dengue vascular leakage is augmented by mast cell degranulation mediated by immunoglobulin Fcgamma receptors. *Elife* 4:e05291. doi: 10.7554/eLife.05291

Takeuchi, O., and Akira, S. (2009). Innate immunity to virus infection. *Immunol. Rev.* 227, 75–86. doi: 10.1111/j.1600-065X.2008.00737.x

Teijaro, J. R., Walsh, K. B., Cahalan, S., Fremgen, D. M., Roberts, E., Scott, F., et al. (2011). Endothelial cells are central orchestrators of cytokine amplification during influenza virus infection. *Cell* 146, 980–991. doi: 10.1016/j.cell.2011.08.015

Thuy, T. B. P., Sugamata, R., Uno, K., Aratani, Y., Ozato, K., Kawachi, S., et al. (2011). Key role of regulated upon activation normal T-cell expressed and secreted, nonstructural protein1 and myeloperoxidase in cytokine storm induced by influenza virus PR-8 (A/H1N1) infection in A549 bronchial epithelial cells. *Microbiol. Immunol.* 55, 874–884. doi: 10.1111/j.1348-0421.2011.00396.x

To, K. F., Chan, P. K. S., Chan, K. F., Lee, W. K., Lam, W. Y., Wong, K. F., et al. (2001). Pathology of fatal human infection associated with avian influenza A H5N1 virus. *J. Med. Virol.* 63, 242–246. doi: 10.1002/1096-9071(200103)63:3<242::AID-JMV1007>3.0.CO;2-N

Uyeki, T. M. (2009). Human infection with highly pathogenic avian influenza A (H5N1) virus: review of clinical issues. *Clin. Infect. Dis.* 49, 279–290. doi: 10.1086/600035

Wong, J. P., Yang, H. M., Nagata, L., Kende, M., Levy, H., Schnell, G., et al. (1999). Liposome-mediated immunotherapy against respiratory influenza virus infection using double-stranded RNA poly ICLC. *Vaccine* 17, 1788–1795. doi: 10.1016/S0264-410X(98)00439-3

Xie, J., Zhang, S., Hu, Y., Li, D., Cui, J., Xue, J., et al. (2014). Regulatory roles of c-jun in H5N1 influenza virus replication and host inflammation. *Biochim. Biophys. Acta* 1842(12 Pt. A), 2479–2488. doi: 10.1016/j.bbadis.2014.04.017

Yu, M., and Levine, S. J. (2011). Toll-like receptor 3, RIG-I-like receptors and the NLRP3 inflammasome: key modulators of innate immune responses to double-stranded RNA viruses. *Cytokine Growth Factor Rev.* 22, 63–72. doi: 10.1016/j.cytogfr.2011.02.001

Zhang, H., Lin, L., Yang, H., Zhang, Z., Yang, X., Zhang, L., et al. (2010). Induction of IL-13 production and upregulation of gene expression of protease activated receptors in P815 cells by IL-6. *Cytokine* 50, 138–145. doi: 10.1016/j.cyto.2010.02.006

Zhang, H., Yang, H., Zhang, L., Yang, X., Zhang, Z., Lin, Q., et al. (2009). Induction of IL-4 release and upregulated expression of protease activated receptors by GM-CSF in P815 cells. *Cytokine* 48, 196–202. doi: 10.1016/j.cyto.2009.07.001

Zhou, J. F., Law, H. K. W., Cheung, C. Y., Ng, I. H. Y., Peiris, J. S. M., and Lau, Y. L. (2006). Differential expression of chemokines and their receptors in adult and neonatal macrophages infected with human or avian influenza viruses. *J. Infect. Dis.* 194, 61–70. doi: 10.1086/504690

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Meng, Huo, Wang, Xiao, Liu, Wei, Dong, Zhang, Hu and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 December 2016
doi: 10.3389/fmicb.2016.02118

Tracking the Evolution of Polymerase Genes of Influenza A Viruses during Interspecies Transmission between Avian and Swine Hosts

*Nipawit Karnbunchob[1], Ryosuke Omori[1,2], Heidi L. Tessmer[1] and Kimihito Ito[1]**

[1] Division of Bioinformatics, Research Center for Zoonosis Control, Hokkaido University, Sapporo, Japan, [2] Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Kawaguchi, Japan

Human influenza pandemics have historically been caused by reassortant influenza A viruses using genes from human and avian viruses. This genetic reassortment between human and avian viruses has been known to occur in swine during viral circulation, as swine are capable of circulating both avian and human viruses. Therefore, avian-to-swine transmission of viruses plays an important role in the emergence of new pandemic strains. The amino acids at several positions on PB2, PB1, and PA are known to determine the host range of influenza A viruses. In this paper, we track viral transmission between avian and swine to investigate the evolution on polymerase genes associated with their hosts. We traced viral transmissions between avian and swine hosts by using nucleotide sequences of avian viruses and swine viruses registered in the NCBI GenBank. Using BLAST and the reciprocal best hits technique, we found 32, 33, and 30 pairs of avian and swine nucleotide sequences that may be associated with avian-to-swine transmissions for PB2, PB1, and PA genes, respectively. Then, we examined the amino acid substitutions involved in these sporadic transmissions. On average, avian-to-swine transmission pairs had 5.47, 3.73, and 5.13 amino acid substitutions on PB2, PB1, and PA, respectively. However, amino acid substitutions were distributed over the positions, and few positions showed common substitutions in the multiple transmission events. Statistical tests on the number of repeated amino acid substitutions suggested that no specific positions on PB2 and PA may be required for avian viruses to infect swine. We also found that avian viruses that transmitted to swine tend to process I478V substitutions on PB2 before interspecies transmission events. Furthermore, most mutations occurred after the interspecies transmissions, possibly due to selective viral adaptation to swine.

Keywords: influenza A viruses, polymerase complex, host range, interspecies transmission, reciprocal best hits

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Masaru Yokoyama,
National Institute of Infectious Diseases, Japan
Hirotaka Ode,
National Hospital Organization Nagoya Medical Center, Japan
Kiyoko Iwatsuki-Horimoto,
University of Tokyo, Japan

****Correspondence:***
Kimihito Ito
itok@czc.hokudai.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *22 October 2016*
Accepted: *15 December 2016*
Published: *26 December 2016*

Citation:
Karnbunchob N, Omori R, Tessmer HL and Ito K (2016) Tracking the Evolution of Polymerase Genes of Influenza A Viruses during Interspecies Transmission between Avian and Swine Hosts. Front. Microbiol. 7:2118. doi: 10.3389/fmicb.2016.02118

INTRODUCTION

The influenza A virus is a negative-sense single-stranded RNA virus that infects humans as well as a wide range of animals (Webster et al., 1992; Kuiken et al., 2004; Tong et al., 2012). Wild aquatic birds, such as wild ducks, geese, gulls, and shorebirds, are the natural reservoirs of the influenza A virus (Kida et al., 1988; Webster et al., 1992). Human influenza pandemics have

historically been caused by genetic reassortment of human and avian influenza A viruses, and this reassortment typically occurs among viruses circulating in swine (Webster and Laver, 1972; Scholtissek et al., 1978; Kawaoka et al., 1989; Yasuda et al., 1991; Smith et al., 2009). Experimental studies have suggested that swine are susceptible to both human (Kundin, 1970) and avian viruses (Kida et al., 1994). Thus, the avian-to-swine transmission of influenza A viruses is an important factor contributing to the emergence of new pandemic strains.

Influenza A viruses are composed of eight gene segments, which encode at least 17 viral proteins (Dubois et al., 2014). Of these, the polymerase complex consisting of PB2, PB1, and PA is responsible for viral replication in host cells. The PB2 protein is responsible for the cap binding of host's mRNA (Webster et al., 1992). The PB1 protein is associated with the catalytic activity of RNA synthesis (Kobayashi et al., 1996; Neumann et al., 2004; Elton et al., 2006). The PA protein is involved in endonuclease activity of the polymerase complex for RNA replication (Dias et al., 2009; Yuan et al., 2009).

The amino acids at several positions on the polymerase complex have been known to determine the host range of influenza A viruses. The amino acid substitution from Glutamic acid (E) to Lysine (K) at position 627 on PB2 of avian viruses increases viral replication in mammalian hosts (Subbarao et al., 1993; Hatta et al., 2001; Shinya et al., 2004; Mok et al., 2014). Two simultaneous amino acid mutations from Valine (V) to Serine (S) at position 715 and from Isoleucine (I) to Serine (S) at position 750 in PB1 are known to reduce the number of cRNA and mRNA (Sugiyama et al., 2009). Several amino acid substitutions in PA were reported to affect viral replication in mammals (Yamayoshi et al., 2014). Most of these studies discuss mammalian adaptation of avian viruses using mouse models of influenza infections. Currently, there is little information about the viral adaptation of avian viruses to swine.

It is important to know which amino acid substitutions on the polymerase complex determine the host range of avian influenza A viruses. A typical alignment-based approach compares consensus sequences of avian viruses and viruses isolated from other hosts, and the different amino acids in their alignments are considered as signature residues for each host (Chen et al., 2006). However, the alignment-based approach is known to be unable to distinguish the founder effect from selective viral adaptation (Tamuri et al., 2009). In order to clarify which amino acid substitutions on viral polymerase are beneficial for avian viruses to transmit to swine, we need to develop a new approach to finding important amino acid substitutions, and each substitution needs to be assessed by statistical tests.

The reciprocal best hits method has been widely used to identify orthologous genes, which are genes shared by different organisms (Tatusov et al., 1997; Bork and Koonin, 1998; Moreno-Hagalsieb and Latimer, 2008). Given two sets of sequences, X and Y, a pair of sequences x in X and y in Y is called a reciprocal best hit, if x is the most similar sequence among X to y and y is the most similar sequence among Y to x. Using a homology search program, such as BLAST, one can retrieve avian virus sequences similar to swine virus sequences. However, if a database contains more than one sequence similar to a sequence associated with a transmission event, simple BLAST searches using a threshold may give multiple combinations of similar sequences. By applying the reciprocal best hits method to the nucleotide sequences of viruses isolated from avian and swine, we can identify pairs of viruses associated with interspecies transmissions without double counting.

In this paper, we investigate the evolution of polymerase genes of influenza A viruses during viral transmission from avian to swine. A pair of nearly identical nucleotide sequences, one of which is from avian viruses and the other from swine, can be considered a footprint of viral transmission between avian and swine hosts. We denote such a pair as a transmission pair. By using BLAST and the reciprocal best hits technique, we explore transmission pairs associated with sporadic transmissions of avian viruses to swine. By analyzing the number of amino acid substitutions on the polymerase proteins found in the transmission pairs of polymerase genes between avian and swine viruses, we examine whether or not these amino acid substitutions are important for interspecies transmission of influenza A viruses between avian and swine hosts.

MATERIALS AND METHODS

Nucleotide Sequences

The nucleotide sequences of PB2, PB1, and PA genes of avian and swine influenza A viruses were downloaded from the National Center for Biotechnology Information (NCBI) Influenza Virus Resource (Bao et al., 2008). Identical nucleotide sequences were removed using the collapse option of the database. Nucleotide sequences containing ambiguous nucleotides or which were less than 95% of the full-length gene were excluded. We obtained 7408, 7531, and 7576 nucleotide sequences of PB2, PB1, and PA genes of influenza A viruses isolated from avian hosts, and 1283, 1340, and 1304 nucleotide sequences of PB2, PB1, and PA genes of influenza A viruses isolated from swine (**Table 1**). We downloaded all the available nucleotide sequences on August 25, 2013.

Bidirectional BLAST Searches between Avian and Swine Viruses

To identify similar nucleotide sequences between swine and avian viruses, we used Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990). For each of the polymerase gene segments,

TABLE 1 | The number of nucleotide sequences of PB2, PB1, and PA genes used in this study.

Hosts	PB2	PB1	PA
Avian	7408	7531	7576
Swine	1283	1340	1304
Total	8691	8871	8880

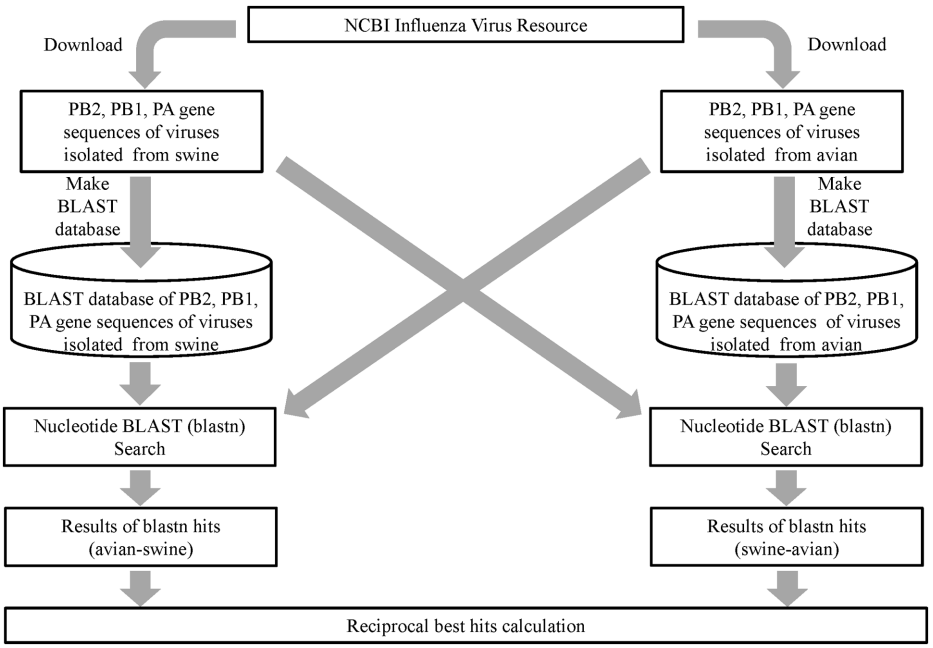


FIGURE 1 | A flow chart for bidirectional BLAST searches between avian and swine influenza A viruses. The PB2, PB1, and PA gene sequences of viruses isolated from avian and swine were downloaded from the NCBI database. For each gene, two BLAST databases were constructed, one from the nucleotide sequences of avian isolates and the other from swine isolates. A blastn search was performed to search for similarities between the avian and swine virus isolates in two directions; first querying avian virus sequences against the swine virus sequence database and vice versa.

we constructed two BLAST databases – one for the nucleotide sequences of avian virus isolates and the other for those of swine virus isolates. BLAST homology searches were conducted bidirectionally using avian sequences as a query against swine sequences as subjects, and vice versa (**Figure 1**). The makeblastdb and blastn commands implemented in ncbi-blast-2.2.28+ were used to construct the databases and to conduct the homology search (Altschul et al., 1990).

Tracking Transmissions by Reciprocal Best Hits Technique

To track the interspecies transmissions of influenza A viruses between avian and swine hosts, we explored avian and swine virus polymerase sequence pairs that are similar to each other by using the reciprocal best hits method. We consider a pair of avian and swine virus sequences that are similar to each other as a footprint of viral transmission between avian and swine hosts, and we call such a pair a transmission pair.

Given two sets of sequences $X = \{x_1, x_2, x_3, \ldots, x_m\}$ and $Y = \{y_1, y_2, y_3, \ldots, y_n\}$, reciprocal best hit pairs can be found as follows: First, for each x_i in X, we perform a BLAST search using x_i against Y and record its top hit as $\mathrm{Top}(x_i)$. Second, for each y_j in Y, we perform a BLAST search using y_j against X and record its top hit as $\mathrm{Top}(y_j)$. Finally, all the pairs of (x_i, y_j) that satisfy $\mathrm{Top}(x_i) = y_j$ and $\mathrm{Top}(y_j) = x_i$ are output as reciprocal best hits.

Figure 2 illustrates how the reciprocal best hits method can track viral transmission events between avian and swine. A1–A6 represent viruses isolated from avian hosts, and S1–S6 represent viruses isolated from swine hosts. Solid lines represent phylogenetic relationships. Dashed arrows represent the top hits found by a blastn search. Pairs (A2, S2) and (A5, S5) are reciprocal best BLAST hits.

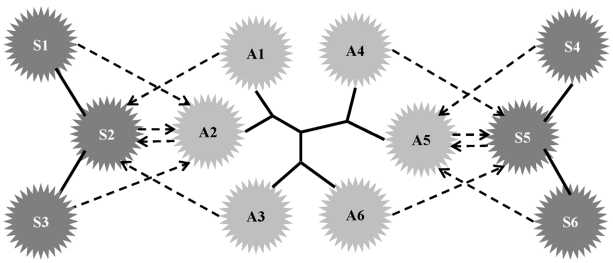


FIGURE 2 | Tracking transmissions by reciprocal best hits. A1–A6 represent viruses isolated from avian hosts. S1–S6 represent viruses isolated from swine hosts. Solid lines represent phylogenetic relationships. Dashed arrows represent the top hits found by a blastn search. A pair of sequences, each of which is the top hit from the other, is called a reciprocal best hit. Pairs (A2, S2) and (A5, S5) are reciprocal best BLAST hits, and we assume such a pair is associated with an interspecies transmission event between avian and swine.

A pair of nucleotide sequences found to be reciprocal best hits and having more than 95% identity with an *E*-value of zero were selected and determined as a transmission pair between avian and swine. We used a custom-made Python program to find reciprocal best hits from BLAST results files. The program is available upon request.

Determination of Transmission Direction by Phylogenetic Analysis

To determine the direction of interspecies transmission, we constructed a phylogenetic tree for each polymerase gene segment. For each polymerase gene, nucleotide sequences of both avian isolates and swine isolates were aligned using Multiple Alignment with Fast Fourier Transform (MAFFT) version 7.245 (Katoh and Standley, 2013). Phylogenetic trees of avian and swine isolates were constructed using the neighbor-joining method (Saitou and Nei, 1987) with ClustalX version 2.1 (Larkin et al., 2007). We used Dendroscope version 3.4.1 (Huson and Scornavacca, 2012) to visualize transmission pairs in phylogenetic trees. A transmission pair found in an avian virus clade was considered an avian-to-swine transmission. In contrast, a transmission pair found in a swine virus clade was considered a possible case of swine-to-avian transmission.

Analysis of Amino Acid Substitutions

To analyze the tendencies in amino acid substitutions in polymerase during avian-to-swine transmission, nucleotide sequences of the transmission pairs of PB2, PB1, and PA genes were translated to protein sequences. For each transmission pair in the avian-to-swine direction, the protein sequence of the avian virus and swine virus were compared, and the amino acid substitutions identified. By the nature of the sequences registered in the database, the positions at the beginning and end were lacking nucleotide information. Gaps found at the beginning and end in transmission pairs were excluded from analysis, and gaps in the other regions were counted in the same way as substitutions.

Statistical Analysis on the Number of Amino Acid Substitutions

If an amino acid position on a polymerase protein determines the host range of viruses, then such a position should be substituted into different amino acids at interspecies transmission events. To determine whether or not some amino acid positions are important for interspecies transmission, we set our null hypothesis to "amino acid substitutions randomly occurred over all positions." We first estimated how many times amino acid substitutions can naturally occur at the same position with random substitutions at independent transmission events.

Let m be the total number of amino acid substitutions occurring on a protein sequence of length l at independent transmission events. Considering multiple transmission events, the total number of amino acid substitutions, m, may exceed the sequence length, l, when we have a large number of transmission events. By assuming amino acid substitutions occur equally over all the positions, the probability that at least one amino acid position is substituted more than n times can be calculated by the following formula:

$$p = 1 - \left(\sum_{k=0}^{n} \binom{m}{k} \left(\frac{1}{l}\right)^{k} \left(1 - \left(\frac{1}{l}\right)\right)^{m-k} \right)^{l} \qquad (1)$$

If the probability for the maximum number of amino acid substitutions in the observed data is smaller than the significance level ($p < 0.05$), we can reject the null hypothesis and conclude that some positions tend to be substituted more frequently than other positions. We confirmed the validity of the formula by comparing it with the multiple substitution probability obtained from Monte Carlo simulations.

Statistical Analysis of Amino Acid Substitutions before and after Avian-to-Swine Transmissions

To characterize the genetic background of avian influenza A viruses that were able to infect swine, we compared consensus amino acid sequences of PB2, PB1, and PA of avian influenza A viruses found in avian-to-swine transmission pairs against consensus amino acids of all avian viruses. Similarly, to characterize the viral adaptation after interspecies transmission from avian to swine, we compared consensus amino acid sequences of PB2, PB1, and PA of swine influenza A viruses found in avian-to-swine transmission pairs against consensus amino acid sequences of all swine viruses. For each position having different consensus amino acids between all avian viruses and avian isolates in avian-to-swine transmission pairs, amino acid variations were further analyzed. We set our null hypothesis to "amino acid compositions at a given position in the two alignments are derived from the same distribution." We use Fisher's exact test (Fisher, 1922) to calculate the probability that the amino acid counts in two alignments come from the same distribution. If this p-value is smaller than the significance level, then the null hypothesis will be rejected.

RESULTS

Transmission Pairs Found in Reciprocal Best Hits

To track the interspecies transmission of influenza A viruses between avian and swine hosts, we looked for nearly identical avian and swine virus polymerase sequences. The reciprocal best hits method found 41, 45, and 45 pairs of avian and swine sequences for PB2, PB1, and PA genes, respectively. All of the reciprocal best hits pairs on the PB2, PB1, and PA genes showed a BLAST E-value of zero. Of these reciprocal best hits pairs, 41 pairs for PB2, 44 pairs for PB1, and 42 pairs for PA had more than 95% identity (Supplementary Tables 1–3). We considered these nearly identical pairs as transmission pairs, which would be associated with transmission of avian viruses to swine or transmission of swine viruses to avian.

The transmission pairs between avian and swine sequences suggested that interspecies transmissions occurred frequently at adjacent places and their isolation years were close to each other. Of 41 transmission pairs for PB2, 32 pairs (78%) were from the same country and 32 pairs (78%) were isolated within 3 years of one another (Supplementary Table 1). Of 44 transmission pairs for PB1, 34 pairs (77%) were from the same country and 33 pairs (75%) were isolated within 3 years (Supplementary Table 2).

Of 42 transmission pairs for PA, 33 pairs (79%) were from the same country and 31 pairs (74%) were isolated within 3 years (Supplementary Table 3). Although there are a few exceptions, these results suggest that transmission occurred between avian and swine located in adjacent areas.

Direction of the Transmission between Avian and Swine

The clade distribution of transmission pairs in phylogenetic trees showed similar trends among PB2, PB1, and PA genes (Supplementary Figure 1). Out of 41 transmission pairs of PB2, 32 (78%) were found in avian clades and 8 (20%) were found in swine clades (Supplementary Table 1). Out of 44 transmission pairs of PB1, 33 (75%) were found in avian clades and 10 (23%) were found in swine clades (Supplementary Table 2). Out of 42 transmission pairs of PA, 30 (71%) were found in avian clades and 11 (26%) were found in swine clades (Supplementary Table 3). A transmission pair found in an avian clade can be considered an avian-to-swine transmission and vice versa. We did not determine the transmission direction for a pair in which one sequence is in an avian clade and the other in a swine clade. In some avian-to-swine transmission pairs, swine viruses were isolated before avian viruses. Similar contrary cases were also observed in swine-to-avian transmissions. The polymerase complex of influenza A viruses is known to evolve slowly because of functional constraints on protein evolution (Gorman et al., 1990). The inconsistency between transmission direction and isolation order may be attributed to the slow evolution of polymerase complex and the delayed viral isolation from their source population. In summary, 78, 75, and 71% of transmission pairs could be associated with avian-to-swine transmission for PB2, PB1, and PA, respectively. In contrast, 20, 23, and 26% of transmission pairs could be swine-to-avian transmissions of PB2, PB1, and PA, respectively.

Amino Acid Substitutions during Avian-to-Swine Transmissions

The PB2 protein is 759 amino acids long, and 175 amino acid substitutions were observed at 142 different positions on PB2 in the 32 avian-to-swine transmission pairs (**Table 2**). An avian-to-swine transmission pair of PB2 has 5.47 amino acid substitutions on average. Note that the count for each position was weighted by the number of transmission pairs having an amino acid substitution at that position, i.e., $(3 \times 5)+(2 \times 23)+(1 \times 114) = 175$, and this total count was averaged by the number of pairs, i.e., $175/32 = 5.47$. When 175 substitutions were randomly distributed over 759 positions, the probability that we observed at least one position substituted four or more times is 0.070 (**Figure 3A**), and the probability that we observed at least one position substituted five or more times is 0.0032 (**Figure 3B**), according to formula (1). To reject the random substitution null hypothesis, we need at least five amino acid substitutions at the same position on the PB2 protein. Among 759 positions on PB2, no position was substituted four or more times. The observed number of multiple amino acid substitutions at the same positions on the PB2 protein was not statistically significant to reject the null hypothesis with a significance level of 0.05. Therefore, we cannot say that avian viruses require amino acid substitutions on specific positions of PB2 to infect swine.

The PB1 protein is 757 amino acids long, and we observed 123 amino acid substitutions at 105 different positions on PB1 in the 33 avian-to-swine transmission pairs (**Table 3**). An avian-to-swine transmission pair on PB1 has 3.73 amino acid substitutions on average. Here, the count for each position was weighted by the number of transmission pairs having an amino acid substitution at that position, i.e., $(4 \times 1)+(3 \times 2)+(2 \times 11)+(1 \times 91) = 123$, and this total count was averaged by the number of pairs, i.e., $123/33 = 3.73$. When 123 substitutions were randomly distributed over 757 positions, the probability that we observed at least one position substituted four or more times is 0.018 (**Figure 3C**), according to formula (1). Among 757 positions on PB1, position 156 was substituted four times. The observed distribution of amino acid substitutions in the PB1 protein was statistically significant to reject the null hypothesis with a significance level of 0.05. However, these four the substitutions were Threonine (T) to Alanine (A), Threonine (T) to Lysine (K), Threonine (T) to Methionine (M), and Methionine (M) to Threonine (T); there was no clear pattern of amino acid substitutions. Furthermore, the consensus amino acid at this position was T in both avian and swine viruses, and the

TABLE 2 | Positions at which amino acid substitutions were observed on PB2 proteins in transmission pairs.

Position at PB2 protein	Number of amino acid substitutions*	Number of positions
389, 473, 547, 570, 674	3	5
8, 129, 139, 147, 194, 255, 327, 338, 344, 354, 355, 446, 456, 461, 575, 613, 661, 667, 684, 699, 701, 756, 759	2	23
15, 23, 28, 51, 58, 64, 66, 67, 79, 81, 87, 89, 92, 95, 105, 106, 109, 113, 122, 130, 134, 136, 137, 142, 144, 152, 153, 157, 164, 168, 175, 177, 184, 186, 187, 188, 189, 191, 192, 199, 206, 221, 240, 253, 265, 274, 285, 292, 293, 295, 299, 300, 303, 307, 309, 312, 313, 315, 317, 318, 330, 336, 340, 343, 351, 352, 356, 365, 368, 374, 375, 376, 394, 395, 399, 416, 422, 433, 444, 451, 463, 464, 468, 480, 488, 503, 508, 509, 521, 526, 537, 543, 553, 574, 584, 588, 598, 618, 629, 648, 659, 663, 666, 678, 680, 702, 714, 715, 717, 723, 731, 732, 742, 748	1	114
Others	0	617
Total		759

The number of amino acid substitutions is defined as the number of transmission pairs having amino acid substitutions at each position. The positions of avian–human signature residues identified by Chen et al. (2006) were underlined.

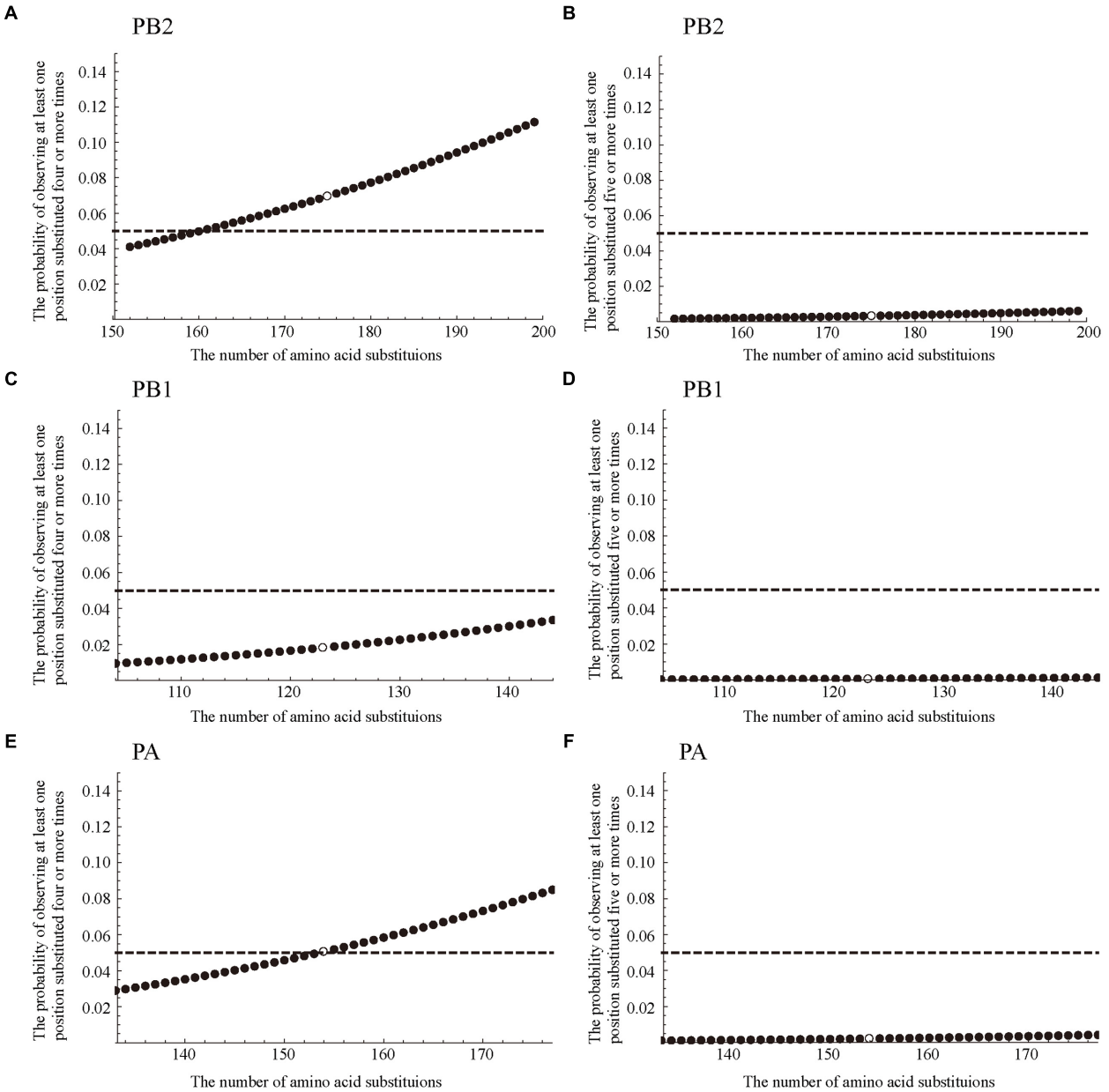


FIGURE 3 | The sensitivity of the significance on the position specific substitution count against the total number of substitutions. White circles show the probability that at least one position is substituted more than given times under the number of substitutions observed in our dataset, and black circles show those of the 95% confidence intervals of the total number of substitutions. **(A,C,E)** show the probabilities of observing at least one position substituted four or more times for PB2, PB1, and PA, respectively. **(B,D,F)** show the probabilities of observing at least one position substituted five or more times for PB2, PB1, and PA, respectively.

distribution of amino acids at position 156 in the transmission pairs showed no clear difference ($p \approx 1.0$ with Fisher's exact test, Supplementary Table 4). There are no research papers showing the importance of this position on the host range, as far as we know. It is difficult to understand a clear reason for such a number of amino acid substitutions at position 156.

The PA protein is 716 amino acids long, and we observed 154 amino acid substitutions at 137 different positions on PA in the 30 avian-to-swine transmission pairs (**Table 4**). An avian-to-swine transmission pair of PA has 5.13 amino acid substitutions on average. Again, the count for each position was weighted by the number of transmission pairs having an amino acid substitution at that position, i.e., $(3 \times 1)+(2 \times 15)+(1 \times 121) = 154$, and this total count was averaged by the number of pairs, i.e., $154/30 = 5.13$. When 154 substitutions were randomly distributed over 716 positions, the probability that we observed

at least one position substituted four or more times is 0.051 (**Figure 3E**), and the probability that we observed at least one position substituted five or more times is 0.0022 (**Figure 3F**), according to formula (1). To reject the random substitution null hypothesis, we need at least five amino acid substitutions at the same position on the PA protein. Among 716 positions on PA, no position was substituted four or more times. The observed distribution of amino acid substitutions in the PA protein was not statistically significant to reject the null hypothesis with a significance level of 0.05. Therefore, we cannot say that avian viruses require amino acid substitutions on specific positions of PA to infect swine.

Analysis of Amino Acid Substitutions before and after Avian-to-Swine Transmissions

To characterize the genetic background of avian influenza A viruses that are able to infect swine, viral adaptation after interspecies transmission from avian to swine hosts was investigated. We compared consensus amino acid sequences of PB2, PB1, and PA for all avian isolates, avian and swine isolates in transmission pairs, and all swine isolates (**Table 5**). Nine positions on PB2, 13 positions on PB1, and five positions on PA had different consensus amino acids when compared to their consensus amino acid sequences. All the positions had the same consensus amino acids between avian and swine isolates on the avian-to-swine transmission pairs. All the positions, except 340 on PB2, had different amino acids between the consensus of swine isolates in transmission pairs and the consensus of all swine isolates, suggesting that positions, except 340 on PB2, were substituted during circulation in swine after avian-to-swine transmission. Amino acids at positions 65, 147, 271, 478, 588, 590, 591, and 645 on PB2, positions 179, 336, 339, 361, 375, 430, 486, 581, 584, 621, 638, 642, and 741 on PB1, and positions 362, 382, 388, 407, and 409 on PA appear to be substituted after interspecies transmission, possibly as a result of selective viral adaptation in swine.

The positions 340 and 478 on PB2 had different amino acids between the consensus of avian isolates in transmission pairs and the consensus of all avian isolates, suggesting that these positions were substituted before avian-to-swine transmission.

The amino acid at position 340 on PB2 of avian viruses in avian-to-swine transmission pairs tended to have Lysine (K), while Arginine (R) in this position is dominant in avian viruses. However, Fisher's exact test on amino acid variation at position 340 showed $p = 0.097$ (**Table 6**), and we cannot reject our null hypothesis that avian PB2 sequences in the transmission pairs

TABLE 3 | Positions at which amino acid substitutions were observed on PB1 proteins in transmission pairs.

Position at PB1 protein	Number of amino acid substitutions*	Number of positions
156	4	1
375, 757	3	2
12, 113, 191, 213, 386, 430, 476, 567, 643, 645, 667	2	11
9, 11, 14, 18, 54, 55, 56, 57, 60, 75, 91, 106, 108, 114, 119, 125, 144, 155, 158, 166, 168, 174, 175, 179, 187, 192, 196, 200, 208, 209, 212, 230, 237, 239, 247, 254, 261, 268, 309, 315, 320, 327, 336, 346, 361, 364, 374, 383, 384, 388, 402, 419, 421, 433, 434, 448, 451, 454, 457, 458, 465, 509, 511, 516, 517, 521, 527, 563, 566, 569, 571, 577, 578, 587, 591, 595, 599, 640, 641, 670, 694, 696, 702, 717, 719, 730, 741, 746, 748, 753, 756	1	91
Others	0	652
Total		757

**The number of amino acid substitutions is defined as the number of transmission pairs having amino acid substitutions at each position. The positions of avian–human signature residues identified by Chen et al. (2006) were underlined.*

TABLE 4 | Positions at which amino acid substitutions were observed on PA proteins in transmission pairs.

Position at PA protein	Number of amino acid substitutions*	Number of positions
441	3	1
57, 61, 80, 94, 186, 208, 231, 261, 272, 388, 554, 560, 585, 626, 684	2	15
3, 12, 14, 22, 27, 29, 38, 40, 78, 86, 88, 89, 90, 101, 109, 115, 118, 127, 135, 141, 158, 159, 169, 192, 201, 204, 210, 216, 224, 226, 238, 247, 252, 256, 262, 269, 277, 285, 308, 311, 312, 315, 316, 321, 323, 325, 335, 336, 345, 350, 353, 355, 361, 364, 372, 394, 403, 407, 413, 418, 424, 435, 437, 439, 440, 442, 450, 451, 469, 487, 490, 492, 496, 499, 501, 505, 512, 513, 518, 522, 528, 531, 535, 536, 538, 545, 547, 553, 558, 565, 566, 569, 570, 576, 580, 581, 587, 588, 589, 592, 595, 598, 602, 607, 615, 616, 617, 618, 619, 621, 625, 628, 665, 669, 680, 683, 693, 694, 696, 706, 715	1	121
Others	0	579
Total		716

**The number of amino acid substitutions is defined as the number of transmission pairs having amino acid substitutions at each position. The positions of avian–human signature residues identified by Chen et al. (2006) were underlined.*

TABLE 5 | Comparison of consensus amino acids on PB2, PB1, and PA among avian isolates, avian and swine isolates in transmission pairs, and swine isolates.

Genes	Position	Consensus amino acid of all avian isolates	Consensus amino acid in avian-to-swine transmission pairs		Consensus amino acid of all swine isolates
			Avian isolates	Swine isolates	
PB2	65	E	E	E	D
	147	I	I	I	T
	271	T	T	T	A
	340	R	K	K	K
	478	I	V	V	I
	588	A	A	A	T
	590	G	G	G	S
	591	Q	Q	Q	R
	645	M	M	M	L
PB1	179	M	M	M	I
	336	V	V	V	I
	339	I	I	I	M
	361	S	S	S	R
	375	N	N	N	S
	430	R	R	R	K
	486	R	R	R	K
	581	E	E	E	D
	584	R	R	R	Q
	621	Q	Q	Q	R
	638	E	E	E	D
	642	N	N	N	S
	741	A	A	A	S
PA	362	K	K	K	R
	382	E	E	E	D
	388	S	S	S	G
	407	I	I	I	V
	409	S	S	S	N

have the same amino acid composition at position 340 with other avian PB2 sequences.

The amino acid at position 478 on PB2 of avian viruses in avian-to-swine transmission pairs tended to have Valine (V), while Isoleucine (I) in this position is dominant in avian viruses. Fisher's exact test on amino acid variation at position 478 showed $p = 6.1 \times 10^{-8}$ (**Table 7**), indicating that avian PB2 sequences in the transmission pairs has a different amino acid composition at position 478 from other avian PB2 sequences. Therefore, I478V mutations on PB2 may be one of the most important amino acid substitutions for avian viruses to transmit to swine.

DISCUSSION

Using BLAST and reciprocal best hits, we found 41, 44, and 42 transmission pairs between avian and swine hosts for PB2, PB1, and PA genes, respectively. These transmission pairs had more than 95% nucleotide identity, indicating that these pairs could be associated with interspecies transmission of influenza A viruses from avian to swine or swine to avian hosts. Phylogenetic analysis showed more than 70% of transmission pairs were associated with avian-to-swine transmissions. By comparing amino acid sequences of avian and swine isolates in the avian-to-swine transmission pairs, we examined amino acid substitutions during avian-to-swine transmissions. On average,

TABLE 6 | Variation of amino acids at position 340 on PB2 of avian viruses.

Amino acid at position 340	Number of sequences	
	Avian isolates excluding avian-to-swine transmission	Avian isolates in avian-to-swine transmission
R	4556	13
K	2797	19
S	12	0
G	8	0
E	2	0
N	1	0
Total	7376	32

The probability that the observed frequencies of amino acids come from the same distribution is 0.097 using Fisher's exact test.

TABLE 7 | Variation of amino acids at position 478 on PB2 of avian viruses.

Amino acid at position 478	Number of sequences	
	Avian isolates excluding avian-to-swine transmission	Avian isolates in avian-to-swine transmission
I	4194	2
V	3147	30
M	30	0
L	4	0
A	1	0
Total	7376	32

The probability that the observed frequencies of amino acids come from the same distribution is 6.1 × 10^{-8} using Fisher's exact test.

avian-to-swine transmission pairs had 5.47, 3.73, and 5.13 amino acid substitutions on PB2, PB1, and PA, respectively. However, amino acid substitutions were distributed over the positions, and few positions showed common substitutions in the multiple transmission events. Statistical tests on the number of repeated amino acid substitutions suggested that no specific positions on PB2 and PA may be required for avian viruses to infect swine. We found that avian viruses involved in avian-to-swine transmissions tended to have Valine (V) at position 478 on PB2, while Isoleucine (I) at position 478 on PB2 are dominant in avian viruses. Statistical tests showed that the distribution of amino acids in avian viruses in avian-to-swine transmissions were different from that of all the avian viruses, suggesting that the I478V substitution may be beneficial for avian viruses to transmit to swine.

Our statistical test is based on the number of amino acid substitutions observed at the same position and the total number of amino acid substitutions at independent transmission events, which are n and m in formula (1), respectively. We assumed the substitution rates among all positions are equivalent with a point estimate of the observed substitution rates. In order to know how the point estimate affects the p-values of statistical tests, we assessed the significance using the 95% confidence intervals (CI) of the total number of amino acid substitutions. From the total number of amino acid substitutions observed in our dataset, the 95% CI of the total number were calculated as [153, 199], [104, 144], and [133, 177] for PB2, PB1, and PA, respectively using the binomial test. Substituting m in formula (1) with numbers in these ranges, we assessed the sensitivity of the significance on the position-specific count of amino acid substitutions to the total number of amino acid substitutions (**Figure 3**). PB2 and PA showed insignificant p-values ($p \geq 0.05$), and we rejected our random null hypothesis for PB2 and PA. However, the significance varied with the total number of amino acid substitutions. PB2 showed insignificant p-values in a wide range of 95% CI on the total number of substitutions. In contrast, PA showed insignificant p-values in half of 95% CI, and the insignificance for PA may be attributed to sampling error. Further data collection is required to assess the significance of the position-specific count of the amino acid substitutions.

Among avian-to-swine transmission pairs of PB2, PB1, and PA genes, some swine viruses possessed different subtypes of HA from their corresponding avian viruses (Supplementary Tables 1–3). These viruses were reassortant viruses receiving HA genes of different subtypes before or after avian-to-swine transmissions. Since the HA protein is associated with receptor specificity in cell entry and is an important determinant of host range, the replacement of the HA subtype may affect amino acid substitutions on the polymerase complex. We examined the effect of HA replacement on the number of amino acid substitutions on PB2, PB1, and PA (Supplementary Tables 5–7). There was no significant difference between transmission pairs with and without HA replacement for PB2 and PB1 ($p \approx 1.0$ for PB2 and PB1). However, transmission pairs of PA had significant differences in the number of amino acid substitutions between pairs having the same HA subtype versus different HA subtypes ($p = 0.043$). Transmission pairs with HA replacement had significantly larger numbers of amino acid substitutions compared to those without HA replacement. Supplementary Tables 8 and 9, respectively, show positions of amino acid substitutions on the avian-to-swine transmission pairs of PA with HA replacement and without HA replacement. The observed numbers of multiple amino acid substitutions at the same positions on the PA were not statistically significant to reject the null hypothesis, when the transmission pairs with and without HA replacement were analyzed separately ($p \geq 0.05$).

The glutamic acid (E) to lysine (K) substitution at position 627 (E627K) on PB2 is known to increase the replication ability of avian influenza viruses in mammalian hosts (Subbarao et al., 1993; Hatta et al., 2001; Shinya et al., 2004; Mok et al., 2014). We did not find this substitution in transmission pairs between avian and swine isolates. **Figure 4** shows three hypotheses that could explain this. Hypothesis A is that the amino acid change at position 627 occurred during the transmission from avian to swine hosts. However, we could not find any instance of this hypothesis. Hypothesis B is that the E627K amino acid change occurred before the transmission from avian to swine hosts, and hypothesis C is that the E627K amino acid changed after the transmission. All of the 32 avian-to-swine transmission pairs possessed E in both avian and swine. Therefore the E627K amino acid substitution on the PB2 protein is not necessary for avian influenza A viruses to infect swine (**Figure 4C**).

Avian viruses involved in avian-to-swine transmissions tended to have R340K and I478V substitutions on PB2. Both positions are known to be residues in the cap-binding domain of PB2 (Guilligay et al., 2008). Although the K at position 340 on PB2 is known to be associated with mammalian adaptation of avian viruses (Xiao et al., 2016), the Fisher's exact test could not reject our null hypothesis. On the other hand, the Fisher's exact test showed a significant difference in amino acid compositions at position 478 on PB2. The I478V substitution may be beneficial for avian viruses to transmit to swine. However, the dominant amino acid at position 478 on PB2 of avian viruses was I, that for avian-to-swine transmission pairs was V, and that for swine viruses was I again (**Table 5**), indicating that it does not determine the host range. It is unclear why position 478 tended to have V only during transmission. Our hypothesis is that the I478V

substitution would be associated with a factor needed for swine to be infected with avian viruses in a natural setting. Experimental studies are needed to determine the effect of this mutation on the tissue tropism, viral growth, polymerase activity, protein expression, and pathogenicity.

Comparing amino acid sequences of influenza A viruses isolated from avian hosts and humans, Chen et al. (2006) identified amino acid positions as signature residues that may be required for avian viruses to infect humans. They have reported 8, 2, and 10 positions of signature residues on PB2, PB1, and PA respectively. Among these positions, amino acid substitutions at four positions (199, 588, 613, and 674) on PB2, two positions (327 and 336) on PB1, and one position (57) on PA were also found in the avian-to-swine transmission pairs in our study (**Tables 2–4**). However, our results suggested that amino acid substitutions at these positions may not be required for avian viruses to infect swine.

Phylogenetic analysis of transmission pairs in reciprocal best hits suggest that interspecies transmissions between avian and swine hosts occur in both directions. Several studies have reported the transmission from avian to swine hosts (Kida et al., 1988; Guan et al., 1996; Karasin et al., 2000; Ninomiya et al., 2002; Choi et al., 2005; Su et al., 2013) and our results on interspecies transmission from avian to swine are consistent with these studies. Experimental research has shown that most avian influenza A virus strains can infect swine (Kida et al., 1994). As described in the results section, avian influenza A viruses may not require specific amino acid substitutions in PB2 and PA to infect swine. Previous studies have also reported phylogenetic evidence of transmission from swine to avian (Olsen et al., 2003; Berhane et al., 2012). Only around 23% of transmission pairs in this study had a swine-to-avian direction. The difference in the number of avian-to-swine and swine-to-avian transmissions may be attributed to the high susceptibility of swine to avian viruses. Another factor that affects the imbalanced transmission direction is the difference in the prevalence of influenza A viruses in avian and swine hosts. The natural reservoirs of the influenza A virus are wild aquatic birds. The prevalence of influenza viruses in the mallard is more than 10% (Olsen et al., 2006), while the prevalence in swine is less than 5% (Corzo et al., 2013). The chance for a pig to be exposed to an avian virus is higher than the chance for a bird to be exposed to a swine virus. Since past pandemics of influenza have been caused by the viral transmission from avian to swine and then swine to human, our result highlights the importance of monitoring avian-to-swine transmission to reduce the chance of future influenza pandemics.

Our reciprocal best hits-based method is applicable to the transmission analysis of other host species or other infectious diseases. In this study, we focused on the interspecies transmission of influenza A viruses between avian and swine hosts. One important future research direction is to analyze transmission of influenza A viruses from avian hosts to other mammalian hosts, including humans, using our method. In our study, we found that avian viruses that transmitted to swine tend to process I478V substitutions on PB2 before interspecies transmission events. By analyzing amino acid substitutions on polymerase during avian-to-human transmissions of H5N1 and H7N9 influenza A viruses, we may be able to identify important amino acid substitutions for avian viruses to transmit to humans. One can also apply the same methodology to analyze the global trend of influenza transmission in humans (Russell et al., 2008). The methodology can also be applied to analyze the transmission of other pathogens, as long as we can access a large amount of their genomic data. Our strategy fully depends on the sequences

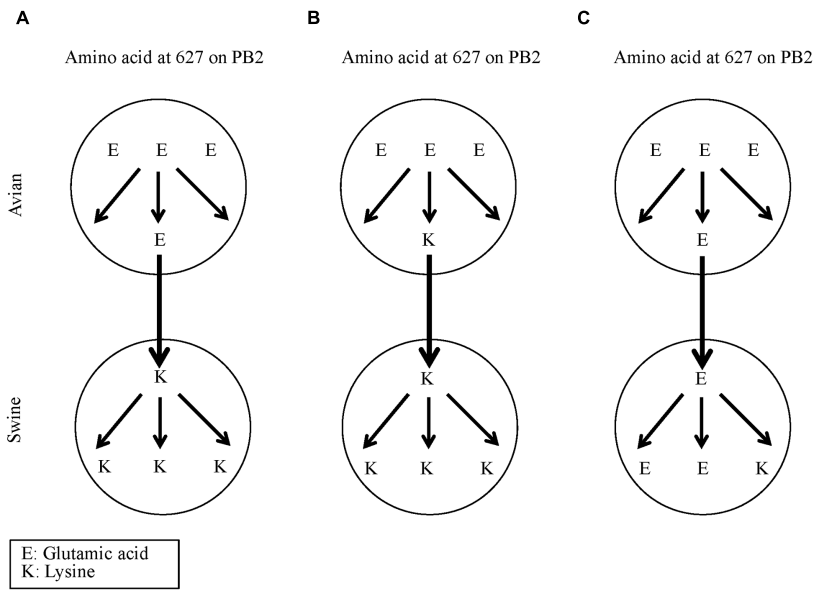


FIGURE 4 | The three possible hypotheses for the E627K substitution on the PB2 protein in the avian-to-swine transmission of influenza A viruses. **(A)** E627K substitution occurred at transmission. **(B)** E627K substitution occurred before transmission. **(C)** E627K substitution occurred after transmission.

registered in the NCBI database. To identify amino acid residues that determine the host range of a virus, we need to assess the importance of amino acid substitutions found in transmission pairs between different host species using a statistical test. If we do not have a sufficient amount of sequence information from a host species, the number of detectable transmission pairs becomes small, and it will be difficult to conduct a statistical test on amino acid substitutions. The greater the quantity of pathogens' nucleotide sequences accumulated in public databases, the higher the chance to obtain meaningful results this method will have.

AUTHOR CONTRIBUTIONS

NK and KI conceived and designed the study. RO and KI designed the statistical analysis. NK analyzed the data. NK, HT, RO, and KI wrote the paper.

FUNDING

This work was supported by CREST (KI) and PRESTO (RO) from Japan Science and Technology Agency (http://www.jst.go.jp/), and by Grant-in-Aid for Scientific Research (B) (KI, 16H02863) and Grant-in-Aid for JSPS Fellows (HT) from the Ministry of Education, Culture, Sports, Science, and Technology in Japan (http://www.mext.go.jp/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.02118/full#supplementary-material

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., et al. (2008). The influenza virus resource at the national center for biotechnology information. *J. Virol.* 82, 596–601. doi: 10.1128/JVI.02005-07

Berhane, Y., Kehler, H., Hande, K., Hisanaga, T., Xu, W., Ojkic, D., et al. (2012). Molecular and antigenic characterization of reassortant H3N2 viruses from turkeys with a unique constellation of pandemic H1N1 internal genes. *PLoS ONE* 7:e32858. doi: 10.1371/journal.pone.0032858

Bork, P., and Koonin, E. V. (1998). Predicting functions from protein sequences–where are the bottlenecks? *Nat. Genet.* 18, 313–318. doi: 10.1038/ng0498-313

Chen, G. W., Chang, S. C., Mok, C. K., Lo, Y. L., Kung, Y. N., Huang, J. H., et al. (2006). Genomic signatures of human versus avian influenza A viruses. *Emerg. Infect. Dis.* 12, 1353–1360. doi: 10.3201/eid1209.060276

Choi, Y. K., Nguyen, T. D., Ozaki, H., Webby, R. J., Puthavathana, P., Buranathal, C., et al. (2005). Studies of H5N1 influenza virus infection of pigs by using viruses isolated in vietnam and thailand in 2004. *J. Virol.* 79, 10821–10825. doi: 10.1128/JVI.79.16.10821-10825.2005

Corzo, C. A., Culhane, M., Juleen, K., Stigger-Rosser, E., Ducatez, M. F., Webby, R. J., et al. (2013). Active surveillance for influenza A virus among swine, midwestern UnitedStates, 2009-2011. *Emerg. Infect. Dis* 19, 954–960. doi: 10.3201/eid1906.121637

Dias, A., Bouvier, D., Crepin, T., McCarthy, A. A., Hart, D. J., Baudin, F., et al. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 458, 914–918. doi: 10.1038/nature07745

Dubois, J., Terrier, O., and Rosa-Calatrava, M. (2014). Influenza viruses and mRNA splicing: doing more with less. *mBio* 5, e00070–14. doi: 10.1128/mBio.00070-14

Elton, D., Digard, P., Tiley, L., and Ortín, J. (2006). "Structure and function of the influenza virus RNP," in *Influenza Virology: Current Topics*, ed. Y. Kawaoka (Norfolk: Caister Academic Press), 1–36.

Fisher, R. A. (1922). On the interpretation of $\chi 2$ from contingency tables, and the calculation of P. *J. R. Stat. Soc.* 85, 87–94. doi: 10.2307/2340521

Gorman, O. T., Donis, R. O., Kawaoka, Y., and Webster, R. G. (1990). Evolution of influenza A virus PB2 genes: implications for evolution of the ribonucleoprotein complex and origin of human influenza A virus. *J. Virol.* 64, 4893–4902.

Guan, Y., Shortridge, K. F., Krauss, S., Li, P. H., Kawaoka, Y., and Webster, R. G. (1996). Emergence of avian H1N1 influenza viruses in pigs in China. *J. Virol.* 70, 8041–8046.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., et al. (2008). The structural basis for cap binding by influenza virus polymerase subunit. *Nat. Struct. Mol. Biol.* 15, 500–506. doi: 10.1038/nsmb.1421

Hatta, M., Gao, P., Halfmann, P., and Kawaoka, Y. (2001). Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* 293, 1840–1842. doi: 10.1126/science.1062882

Huson, D. H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 6, 1061–1067. doi: 10.1093/sysbio/sys062

Karasin, A. I., Brown, I. H., Carman, S., and Olsen, C. W. (2000). Isolation and characterization of H4N6 avian influenza viruses from pigs with pneumonia in Canada. *J. Virol.* 74, 9322–9327. doi: 10.1128/JVI.74.19.9322-9327.2000

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kawaoka, Y., Krauss, S., and Webster, R. G. (1989). Avian-to-human transmission of the PB1 gene of influenza A virus. *J. Virol.* 63, 4603–4608.

Kida, H., Ito, T., Yasuda, J., Shimizu, Y., Itakura, C., Shortridge, K. F., et al. (1994). Potential for transmission of avian influenza viruses to pigs. *J. Gen. Virol.* 75, 2183–2188. doi: 10.1099/0022-1317-75-9-2183

Kida, H., Shortridge, K. F., and Webster, R. G. (1988). Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs in China. *Virology* 162, 160–166. doi: 10.1016/0042-6822(88)90405-9

Kobayashi, M., Toyoda, T., and Ishihama, A. (1996). Influenza virus PB1 protein is the minimal and essential subunit of RNA polymerase. *Arch. Virol* 141, 525–539. doi: 10.1007/BF01718315

Kuiken, T., Rimmelzwaan, G., van Riel, D., van Amerongen, G., Baars, M., and Fouchier, R. (2004). Avian H5N1 influenza in cats. *Science* 306, 241. doi: 10.1126/science.1102287

Kundin, W. D. (1970). Hong Kong A-2 influenza virus infection among swine during a human epidemic in Taiwan. *Nature* 228, 857. doi: 10.1038/228857a0

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Mok, C. K., Lee, H. H., Lestra, M., Nicholls, J. M., Chan, M. C., Sia, S. F., et al. (2014). Amino acid substitutions in polymerase basic protein 2 gene contribute to the pathogenicity of the novel A/H7N9 influenza virus in mammalian hosts. *J. Virol.* 88, 3568–3576. doi: 10.1128/JVI.02740-13

Moreno-Hagalsieb, G., and Latimer, K. (2008). Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319–324. doi: 10.1093/bioinformatics/btm585

Neumann, G., Brownlee, G. G., Fodor, E., and Kawaoka, Y. (2004). Orthomyxovirus replication, transcription, and polyadenylation. *Curr. Top. Microbiol. Immunol.* 283, 121–143.

Ninomiya, A., Takada, A., Okazaki, K., Shortridge, K. F., and Kida, H. (2002). Seroepidemiological evidence of avian H4, H5, and H9 influenza A virus transmission to pigs in southeastern China. *Vet. Microbiol.* 88, 107–114. doi: 10.1016/S0378-1135(02)00105-0

Olsen, B., Munster, V. J., Wallensten, A., Waldenström, J., Osterhaus, A. D., and Fouchier, R. A. (2006). Global patterns of influenza a virus in wild birds. *Science* 312, 384–388. doi: 10.1126/science.1122438

Olsen, C. W., Karasin, A., and Erickson, G. (2003). Characterization of a swine-like reassortant H1N2 influenza virus isolated from a wild duck in the United States. *Virus Res.* 93, 115–121. doi: 10.1016/S0168-1702(03)00073-X

Russell, C. A., Jones, T. C., Barr, I. G., Cox, N. J., Garten, R. J., Gregory, V., et al. (2008). The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320, 340–346. doi: 10.1126/science.1154137

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Scholtissek, C., Rohde, W., Von Hoyningen, V., and Rott, R. (1978). On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* 87, 13–20. doi: 10.1016/0042-6822(78)90153-8

Shinya, K., Hamm, S., Hatta, M., Ito, H., Ito, T., and Kawaoka, Y. (2004). PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* 320, 258–266. doi: 10.1016/j.virol.2003.11.030

Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459, 1122–1125. doi: 10.1038/nature08182

Su, S., Qi, W., Chen, J., Zhu, W., Huang, Z., Xie, J., et al. (2013). Seroepidemiological evidence of avian influenza A Virus transmission to pigs in southern china. *J. Clin. Microbiol.* 51, 601–602. doi: 10.1128/JCM.02625-12

Subbarao, E. K., London, W., and Murphy, B. R. (1993). A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* 67, 1761–1764.

Sugiyama, K., Obayashi, E., Kawaguchi, A., Suzuki, Y., Tame, J. R., Nagata, K., et al. (2009). Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *EMBO J.* 28, 1803–1811. doi: 10.1038/emboj.2009.138

Tamuri, A. U., dos Reis, M., Hay, A. J., and Goldstein, R. A. (2009). Identifying changes in selective constraints: host shifts in influenza. *PLoS Comput. Biol.* 5:e1000564. doi: 10.1371/journal.pcbi.1000564

Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997). A genomic perspective on protein families. *Science* 278, 631–637. doi: 10.1126/science.278.5338.631

Tong, S., Li, Y., Rivailler, P., Conrardy, C., Castillo, D. A., Chen, L. M., et al. (2012). A distinct lineage of influenza A virus from bats. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4269–4274. doi: 10.1016/S0378-1135(02)00105-0

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56, 152–179.

Webster, R. G., and Laver, W. G. (1972). Studies on the origin of pandemic influenza. 1. Antigenic analysis of A2 influenza viruses isolated before and after the appearance of Hong Kong influenza using antisera to the isolated hemagglutinin subunits. *Virology* 48, 433–444. doi: 10.1016/0042-6822(72)90054-2

Xiao, C., Ma, W., Sun, N., Huang, L., Li, Y., Zeng, Z., et al. (2016). PB2-588 V promotes the mammalian adaptation of H10N8, H7N9 and H9N2 avian influenza viruses. *Sci. Rep.* 6, 19474. doi: 10.1038/srep19474

Yamayoshi, S., Yamada, S., Fukuyama, S., Murakami, S., Zhao, D., Uraki, R., et al. (2014). Virulence-affecting amino acid changes in the PA protein of H7N9 influenza A viruses. *J. Virol.* 88, 3127–3134. doi: 10.1128/JVI.03155-13

Yasuda, J., Shortridge, K. F., Shimizu, Y., and Kida, H. (1991). Molecular evidence for a role of domestic ducks in the introduction of avian H3 influenza viruses to pigs in southern China, where the A/Hong Kong/68 (H3N2) strain emerged. *J. Gen. Virol.* 72, 2007–2010. doi: 10.1099/0022-1317-72-8-2007

Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., et al. (2009). Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* 458, 909–913. doi: 10.1038/nature07720

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Karnbunchob, Omori, Tessmer and Ito. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 July 2016
doi: 10.3389/fmicb.2016.01068

A Novel H1N2 Influenza Virus Related to the Classical and Human Influenza Viruses from Pigs in Southern China

Yafen Song[1,2,3,4†], *Xiaowei Wu*[1,5†], *Nianchen Wang*[1,2,3,4], *Guowen Ouyang*[1,2,3,4], *Nannan Qu*[1,2,3,4], *Jin Cui*[1,2,3,4], *Yan Qi*[6], *Ming Liao*[1,2,3,4*] *and Peirong Jiao*[1,2,3,4*]

[1] College of Veterinary Medicine, South China Agricultural University, Guangzhou, China, [2] National and Regional Joint Engineering Laboratory for Medicament of Zoonosis Prevention and Control, Guangzhou, China, [3] Key Laboratory of Animal Vaccine Development, Ministry of Agriculture, Guangzhou, China, [4] Key Laboratory of Zoonosis Prevention and Control of Guangdong, Guangzhou, China, [5] Guangdong Entry-Exit Inspection and Quarantine Bureau, Guangzhou, China, [6] China Animal Husbandry Group, Beijing, China

Southern China has long been considered to be an epicenter of pandemic influenza viruses. The special environment, breeding mode, and lifestyle in southern China provides more chances for wild aquatic birds, domestic poultry, pigs, and humans to be in contact. This creates the opportunity for interspecies transmission and generation of new influenza viruses. In this study, we reported a novel reassortant H1N2 influenza virus from pigs in southern China. According to the phylogenetic trees and homology of the nucleotide sequence, the virus was confirmed to be a novel triple-reassortant H1N2 virus containing genes from classical swine (PB2, PB1, HA, NP, and NS genes), triple-reassortant swine (PA and M genes), and recent human (NA gene) lineages. It indicated that the novel reassortment virus among human and swine influenza viruses occurred in pigs in southern China. The isolation of the novel reassortant H1N2 influenza viruses provides further evidence that pigs are "mixing vessels," and swine influenza virus surveillance in southern China will provide important information about genetic evaluation and antigenic variation of swine influenza virus to formulate the prevention and control measures for the viruses.

Keywords: swine influenza virus, H1N2, reassortant, phylogenetic analysis, molecular characterization

OPEN ACCESS

Edited by:
José A. Melero,
Instituto de Salud Carlos III, Spain

Reviewed by:
Julie McAuley,
The Peter Doherty Institute for Infection and Immunity, Australia
Shailesh D. Pawar,
National Institute of Virology, India

***Correspondence:**
Ming Liao
mliao@scau.edu.cn
Peirong Jiao
prjiao@scau.edu.cn

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *14 April 2016*
Accepted: *24 June 2016*
Published: *08 July 2016*

Citation:
Song Y, Wu X, Wang N, Ouyang G, Qu N, Cui J, Qi Y, Liao M and Jiao P (2016) A Novel H1N2 Influenza Virus Related to the Classical and Human Influenza Viruses from Pigs in Southern China. Front. Microbiol. 7:1068. doi: 10.3389/fmicb.2016.01068

INTRODUCTION

China is the biggest country for swine breeding and pork production, and the largest market for the consumption of pork in the world. It is also the only region that frequently imports pigs from other continents (Zhu et al., 2013; Kong et al., 2014). Swine influenza is an acute respiratory viral disease characterized by coughing, sneezing, nasal discharge, elevated rectal temperatures, lethargy, difficult breathing, and depressed appetite that decreases health and welfare of pigs and results in a significant economic loss for the swine industry worldwide (Kothalawala et al., 2006). As the RNA-dependent RNA polymerase could not proofread the newly synthesized gene segments, mutations of influenza A viruses arise in each replication cycle (Nichol et al., 2000; Koçer et al., 2013). It comes as no surprise; the viruses never stop changing and generating new viral subtypes via mutation, recombination and reassortant. Currently, H1N1, H3N2, and H1N2 swine influenza viruses (SIVs) are mainly circulating in the swine population in China, but H3N8, H4N8, H5N1, H6N6, and H9N2 influenza viruses have been also isolated in swine in China (Kong et al., 2014). Overall,

special breeding environments and features of the virus result in various subtypes and genotype coexisting in China, which would accelerate the genomics evolutionary and antigenic variation of swine influenza viruses. Therefore, virologic and serologic surveillance of swine influenza virus is urgently required.

From January 2012 to March 2012, we carried out swine influenza virus surveillance in southern China. We collected 300 samples from pigs on swine farms in Guangdong province. One swine influenza virus was isolated. We analyzed the origin, genetic composition, and antigenicity characteristics of the hemagglutinin (HA) protein of the novel H1N2 subtypes isolated from pigs.

MATERIALS AND METHODS

Sample Collection

From January 2012 to March 2012, we monitored swine influenza virus in swine farms in Guangdong province of southern China. We chose 30 swine farms and randomly collected 300 nasal swab samples from the 5 to 9-month-old fattening pigs, and also from sows, weaning pigs, nursery pigs, and boars which showed suspicious clinical symptoms. The samples were sent to our laboratory, and stored at −80°C until analysis. This study of nasal sampling from pigs was carried out in accordance with the recommendations of the experimental animal administration and ethics committee of South China Agriculture University of guidelines. The protocol was approved by the biosafety committee of South China Agriculture University.

Virus Isolation and Identification

The collected samples were inoculated into amnionic and allantoic cavities of 9–10-day-old specific-pathogen-free (SPF) embryonated chickens eggs. After incubating for 48 h at 37°C, the allantoic fluids were harvested, and the reverse-transcription polymerase-chain reaction (RT-PCR), hemagglutination test, and hemagglutination inhibition (HI) test were performed to identify and subtype the positive influenza samples as described previously (Song et al., 2015). Finally, virus allantoic fluids were harvested and stored at −80°C before use. All experiments were carried out in ABSL-3 facilities in compliance with the biosafety committee of South China Agriculture University protocols. All experiments handling was performed in accordance with the experimental animal administration and ethics committee of South China Agriculture University guidelines.

Gene Sequence and Molecular Analysis

Viral RNA was extracted from allantoic fluid with Trizol LS Reagent (Life Technologies, Inc.) and transcribed into cDNA with SuperScript III reverse transcriptase (Invitrogen, China). PCR was performed as described previously (Song et al., 2015). The products were purified with the QIAquick PCR purification kit (QIAGEN) following the manufacturer's instructions and sequencing was performed by using an ABI Prism 3730 genetic analyzer (Applied Biosystems) by Shanghai Invitrogen Biotechnology Co., Ltd.

Sequencing data were compiled with the SEQMAN program of Lasergene 7 (DNASTAR). All the referred sequences of this article were downloaded from NCBI databases. BLAST analysis was carried out on NCBI. The consensus sequences of each lineage were obtained using MegAlign and then compared with MEGA (version 4.0) using Clustal W Method. Phylogenetic trees were generated with MEGA program (version 4.0) using neighbor-joining analysis. Bootstrap value was calculated on 1000 replicates of the alignment. The nucleotide sequences in this study are available on GenBank (accession number KX269879-KX269886).

RESULTS

Virus Isolation and Identification

The virus was isolated from nasal swab samples of nursery pigs, and identified by the reverse-transcription polymerase-chain reaction (RT-PCR), hemagglutination test, and hemagglutination inhibition (HI) test and confirmed by genomic sequencing and the nucleotide BLASTn analysis. According to the results, the virus was identified as swine influenza A (H1N2) virus and named as A/swine/Guangdong/1/2012(H1N2).

Homology Analysis of Nucleotide Sequences

To understand whether the swine H1N2 isolate [A/swine/Guangdong/1/2012(H1N2)] is related to the previous swine H1N2 viruses, eight gene segments of the virus were sequenced, and the homology was determined by comparison with the sequences available in GenBank. Viral homology analysis of nucleotide sequences of the virus was presented in **Table 1** and **Figure 9**. The PB2 gene of the virus shared the highest nucleotide sequence identity with a triple-reassortant swine (TRIG) H1N2 influenza virus (Vijaykrishna et al., 2011), A/swine/Hong Kong/NS30/2004 (H1N2), with a homology rate of 97.5%. The PB1 shared the highest nucleotide sequence identity with another TRIG H1N2 influenza virus [A/swine/Hong Kong/NS1890/2009 (H1N2)], with a homology rate 96.9%. The PA and NP genes both shared the highest nucleotide sequence identity with

TABLE 1 | Nucleotide sequence identity between the isolate virus and reference strains available in GenBank[a].

Gene[b]	Virus with the greatest similarity	Identity (%)
HA	A/swine/Hainan/1/2005(H1N2)	96.8
NA	A/swine/Hong Kong/294/2009(H1N2)	97.1
PB2	A/swine/Hong Kong/NS30/2004(H1N2)	97.5
PB1	A/swine/Hong Kong/NS1890/2009(H1N2)	96.9
PA	A/swine/Hong Kong/1111/2004(H1N2)	97.6
NP	A/swine/Hong Kong/1111/2004(H1N2)	97.1
M	A/swine/Hong Kong/NS623/2002(H1N2)	97.9
NS	A/swine/Hong Kong/NS623/2002(H1N2)	97.1

[a] *http://www.ncbi.nlm.nih.gov/.*

[b] *HA, hemagglutinin; NA, neuraminidase; PB, polymerase basic subunit; PA, polymerase acidic subunit; NP, nucleoprotein; M, matrix; NS, nonstructural.*

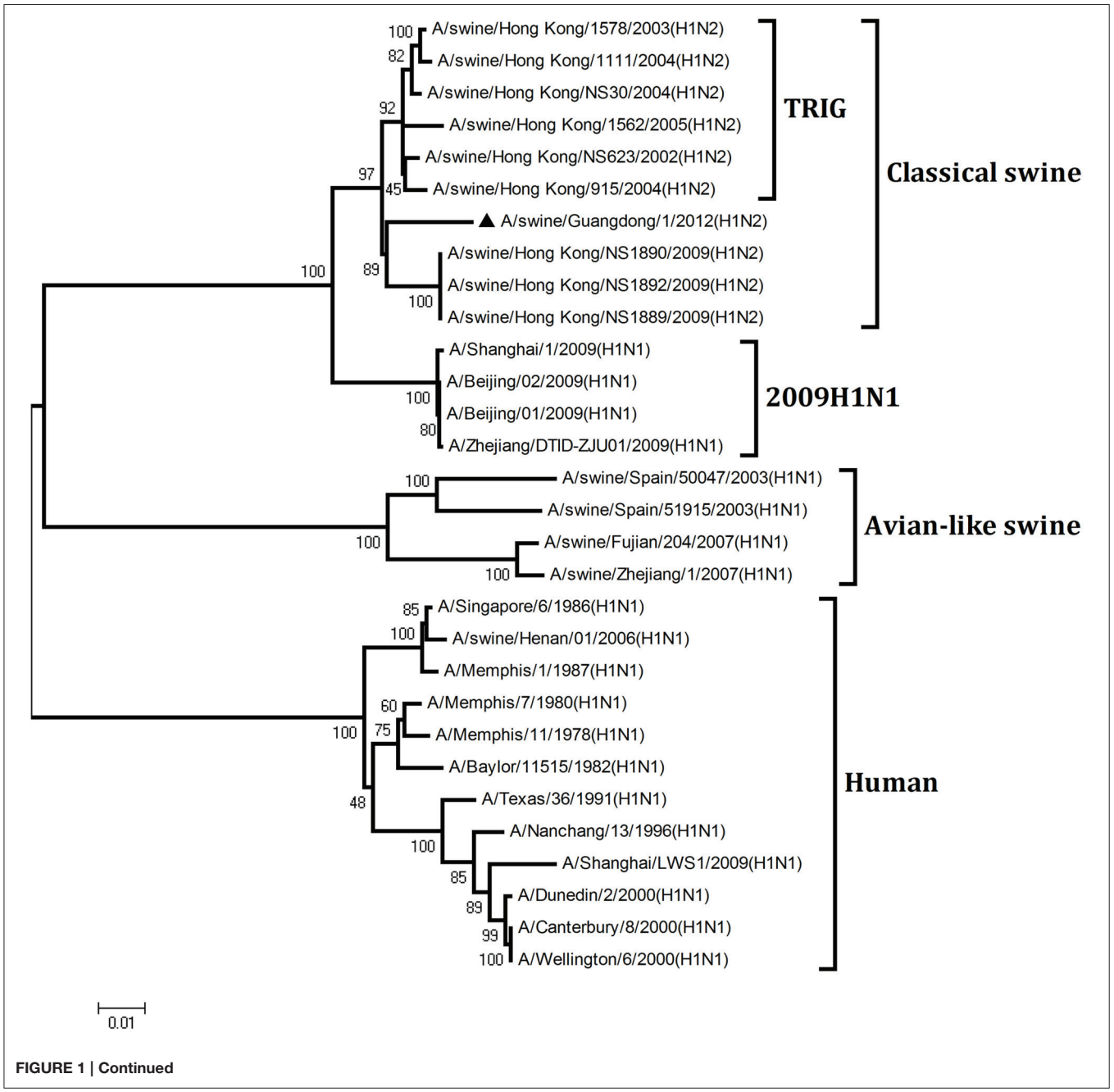


FIGURE 1 | Continued

A/swine/Hong Kong/1111/2004 (H1N2), which was another TRIG H1N2 influenza virus isolated in Hong Kong, with a homology rate ranging from 97.1 to 97.6%. The HA gene was the most closely related to the classical swine (CS) H1N2 viruses (A/swine/Hainan/1/2005 (H1N2; Yu et al., 2009), with a homology rate 96.8%. The NA gene showed a close relationship with A/swine/Hong Kong/294/2009 (H1N2; with a homology rate 97.1%), which was a TRIG H1N2 influenza virus that acquired the HA gene from the CS viruses (Smith et al., 2009). The M and NS genes shared the highest nucleotide sequence identity with another TRIG H1N2 influenza virus (Smith et al., 2009), A/swine/Hong Kong/NS623/2002 (H1N2), with a homology rate ranging from 97.1 to 97.9%. Thus, the results of the homology analysis suggested that the virus might be a multi-reassortant virus.

Phylogenetic Analysis of the Virus

To understand the genetic origin of the gene segments of the virus more precisely, eight phylogenetic trees were constructed using the nucleotide sequences of the virus and the genes of reference viruses available in GenBank, which included viruses isolated from poultry, human, and swine.

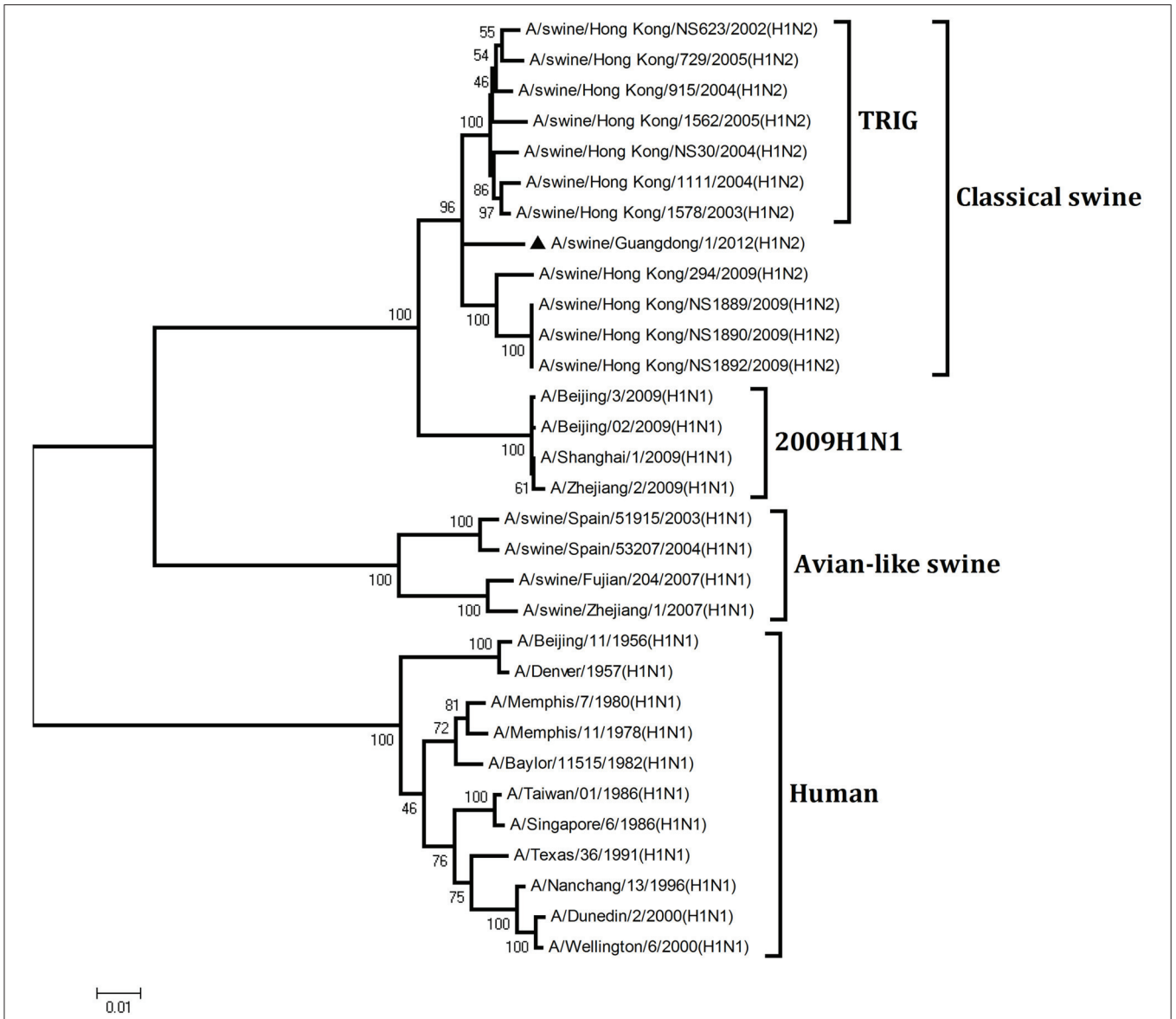


FIGURE 2 | Continued

The phylogenetic analysis results showed that the PB2, PB1, HA, NP, and NS genes of the novel H1N2 virus all fell into the classical swine lineage (**Figures 1**, **2**, **4**, **5**, **8**). The PA and M genes of the virus belonged to the TRIG lineage (**Figures 3**, **7**). The NA gene of A/swine/Guangdong/1/2012(H1N2) was special and segregated into recent human lineage, early human lineage, earliest human lineage, and avian-like swine lineage (**Figure 6**). Though the NA gene of the virus was closely related to that of those TRIG H1N2 influenza viruses, it fell into the recent human lineage, which containing human H3N2 influenza viruses isolated from the 1990s and the Twenty-first century. Therefore, according to the phylogenetic trees and homology of the nucleotide sequence, the virus was confirmed to be a novel triple-reassortant H1N2 virus containing genes from classical swine (PB2, PB1, HA, NP, and NS genes), TRIG (PA and M genes), and recent human (NA gene) lineages.

Molecular Analysis

In our study, the deduced amino acid sequences of the HA region of the novel H1N2 virus and other representative influenza viruses from China were aligned and analyzed. The novel H1N2 virus and other representative viruses all contained an amino acid motif PSIQSR↓G at their HA cleavage sites, which is a characteristic of low pathogenic influenza viruses. Five potential glycosylation sites (N–X–S/T) were conserved at positions 27, 28, 40, 498, and 557 (H1 numbering) in the HA protein of all analyzed viruses (**Figure 10**). The viruses from human, classical

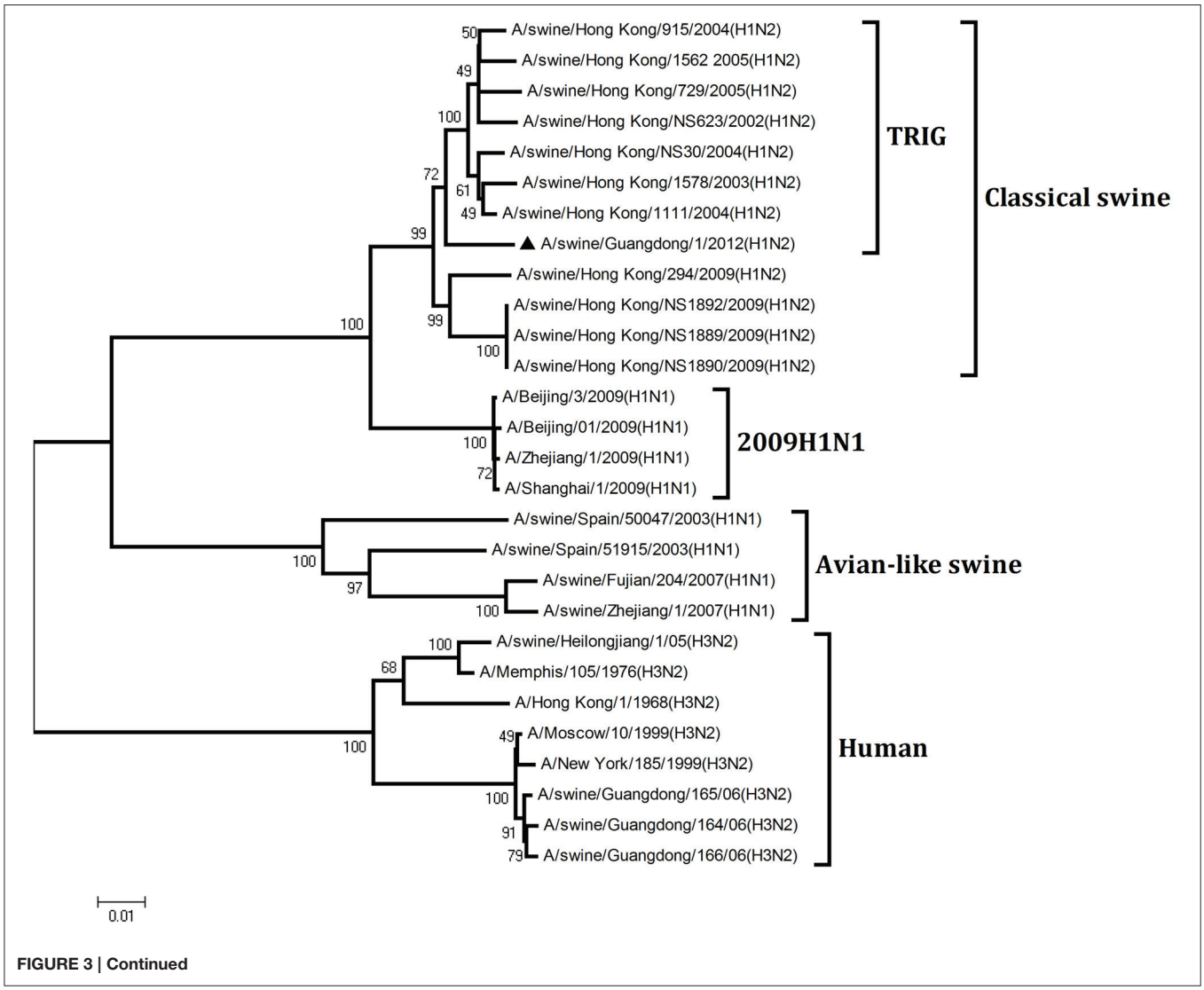


FIGURE 3 | Continued

swine, and 2009H1N1 lineages also had conserved glycosylation sites at positions 104 and 304 (H1 numbering). One potential glycosylation sites at position 293 (H1 numbering) only existed in some viruses. The K136N (H1 numbering) mutation brought a new potential glycosylation site at position 136 in the HA protein of the novel H1N2 virus. In addition, the viruses of the human, classical swine, and 2009H1N1 lineages had more glycosylation sites than those of the avian-like lineage. The novel H1N2 virus had the same D at positions 225 of HA as the 1918 human strain, suggesting that the virus may have potential to infect humans. Antigenic sites are regions of molecules involved in antibody binding. Several mutations were found in the antigenic sites of the A/swine/Guangdong/1/2012(H1N2) virus (**Figure 10**): G158E at site Sa; S188N, A189I, and D204E at site Sb; S140P, H141Y, N145R, I169L, S206T, K224R, and Q240G at site Ca; and S74F and N77S at site Cb (H3 numbering).

In our study, we found five glycosylation sites in the NA protein in the novel H1N2 virus. Two were in the linker region (N61 and N70) and the other three were in the NA domain (N146, N200, and N234). The amino acid substitutions (E119G, H274Y, R292K, and N294S) were not observed in the NA protein of the novel H1N2 virus, which suggested the virus was still sensitive to NA inhibitors.

Analysis based on the deduced amino acid sequence of A/swine/Guangdong/1/2012(H1N2) and its potential donor virus, the novel H1N2 virus and its potential donor virus [A/swine/Hong Kong/NS30/2004 (H1N2)] both contained 271A, 590S, 591R, 627E, and 701D in the PB2 protein. The isolate and its potential donor virus [A/Hong Kong/NS1890/2009(H1N2)] both contained a truncated PB1-F2 protein (57aa) and a PB1-N40 protein (718aa). The PA-X protein of swine influenza viruses usually exists in either full length or a truncated form (either 61aa or 41aa). The novel H1N2 virus and A/Hong Kong/1111/2004 (H1N2) both possessed a full-length PA-X. The amino acid substitutions (26, 30, 31, and 34) were not observed in the M2 protein of the novel H1N2 virus and its potential donor virus

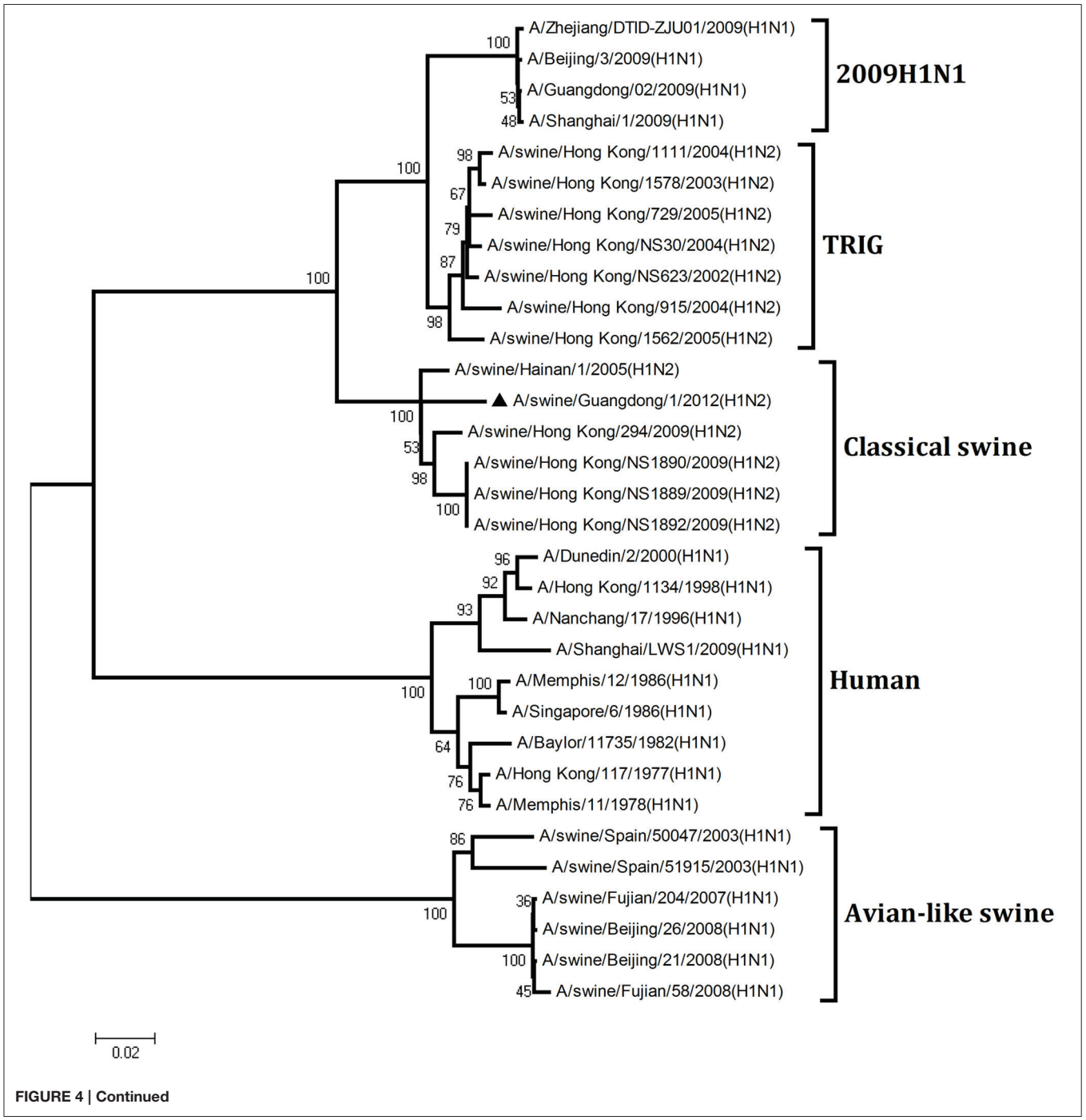


FIGURE 4 | Continued

A/swine/Hong Kong/NS623/2002(H1N2). But V27I substitution both occurred in the M2 protein of them.

DISCUSSION

Swine influenza was first observed as a pertinent disease of swine in 1918 at the time of the human pandemic, and the virus was isolated and identified in 1930. This is known as the "classical" H1N1 swine virus. From 1918 to 1919, the classical swine influenza virus caused a high mortality among pigs in Chinese coastal cities. The classical swine virus was first isolated in Hong Kong of China in 1974 and continued to presence in apparently healthy pigs in Hong Kong and Mainland China (Zhu et al., 2013). Recently, classical H1N1 swine virus emerged in humans as a reassortant (2009/H1N1) and caused the 2009 H1N1 influenza pandemic (Dawood et al., 2009). Reassortant H1N2 influenza A viruses derived from human-like swine H3N2 and classical swine H1N1 viruses were first isolated in Japan

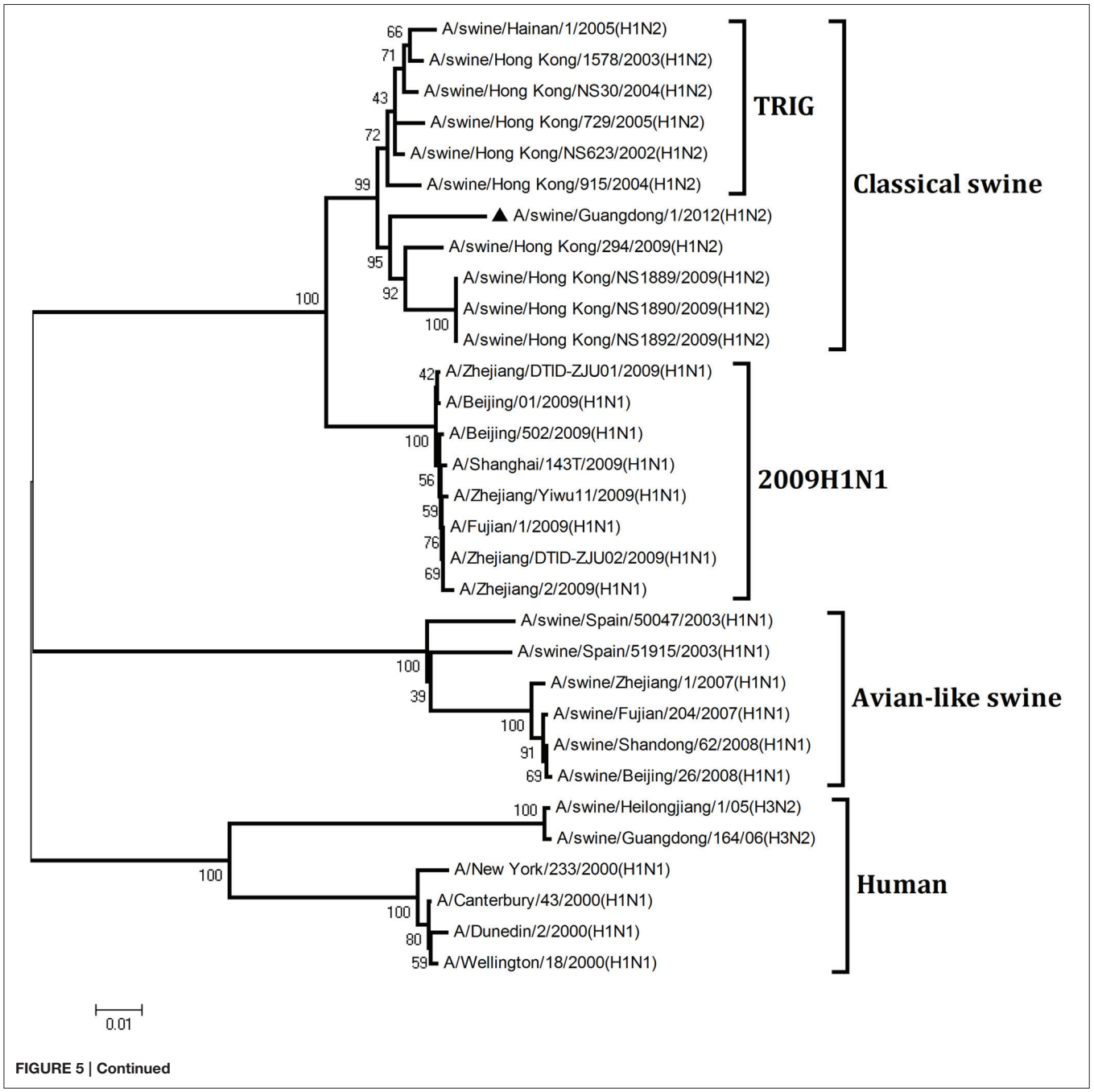


FIGURE 5 | Continued

in 1978 and became endemic in Japanese swine populations (Sugimura et al., 1980). Then, these reassortant swine H1N2 influenza viruses have also been isolated and demonstrated in many countries, such as France, the United Kingdom, the United States, Korea, Spain, Germany, Thailand, and so on (Yu et al., 2009). However, these reassortant H1N2 viruses were not reported until 2004 in China (Qi and Lu, 2006).

In our study, a reassortant swine influenza virus (A/swine/Guangdong/1/2012) was isolated from nasal swab samples of nursery pigs when we collected a total of 300 samples from pigs in swine farms of Guangdong province during swine influenza virus surveillance between January 2012 and March 2012. Homology analysis showed that the HA gene was most closely related to the classical swine (CS) H1N2 viruses and the NA gene showed a closer relationship with a TRIG H1N2 influenza virus that acquired the HA gene from the CS viruses. Phylogenetic analysis showed that the PB2, PB1, HA, NP, and NS genes of the virus fell into the classical swine lineage, and the NA gene belonged to the recent human lineage. The PA and M genes of the virus belonged to TRIG lineage. The

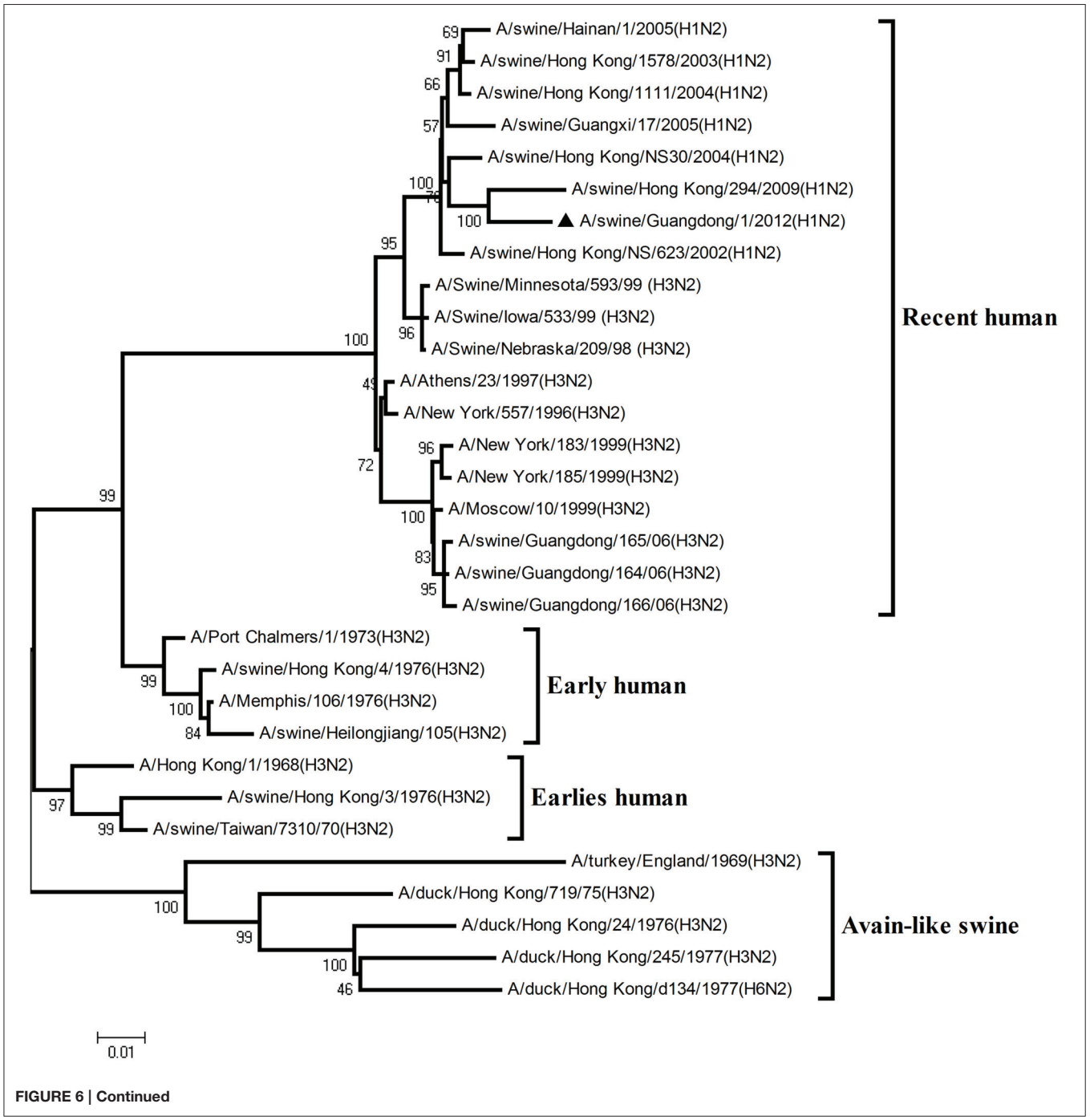


FIGURE 6 | Continued

data confirmed that A/swine/Guangdong/1/2012(H1N1) was a novel triple-reassortant H1N2 virus, and the recombination event occurred in swine populations of southern China.

In general, species barriers prevent the movement of influenza viruses from one host into another host. For complete adaptation and further transmission between species, influenza A virus must overcome species barriers including spatial, physiological, and molecular in origin (Koçer et al., 2013). When influenza A virus shed by one host to infect another, it must breach entry barriers. The specificity of receptor molecules usually governs viral entry into cells. It is well-known that avian influenza viruses preferentially bind to sialic acids (SA)-α-2,3-Gal–terminated glycoproteins, whereas human influenza viruses bind SA-α-2,6-Gal–terminated glycoproteins. In humans, epithelial cells of the upper airway predominantly express SA-α-2,6-Gal–terminated glycoproteins. These differences may be responsible for restricting replication of avian viruses

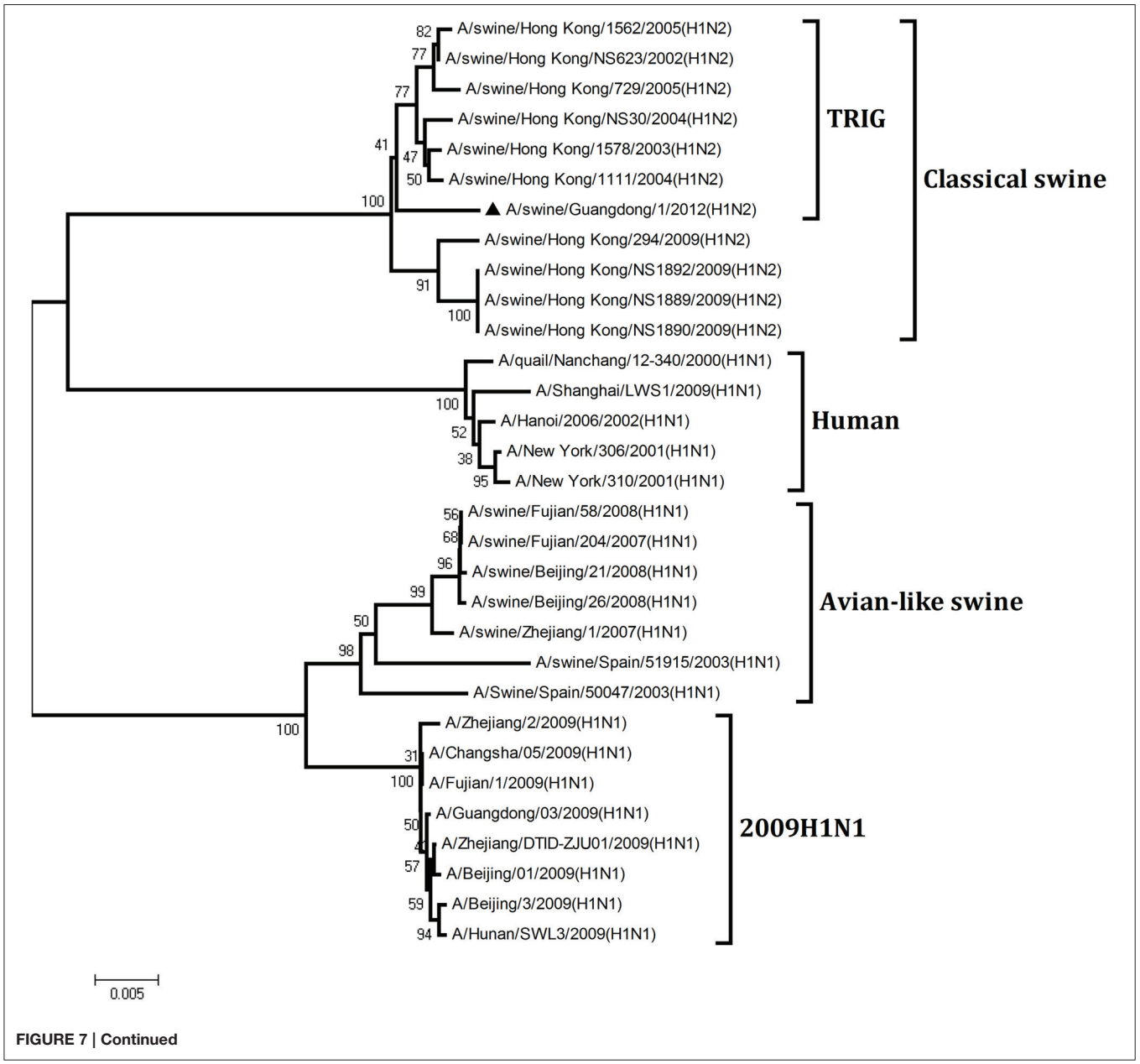


FIGURE 7 | Continued

in humans. However, respiratory epithelia of pigs express both SA-α-2,3-Gal–terminated glycoproteins and SA-α-2,6-Gal–terminated glycoproteins. Thus, pigs are susceptible to infection with both avian and human influenza virus suggesting that pigs are a potential source for generating novel reassortant influenza viruses and are more frequently involved in interspecies transmission of influenza A viruses than other animals. Therefore, pigs are always considered "mixing vessels" (Brown, 2001; Kuiken et al., 2006). The outbreak of the swine-origin triple-reassortant H1N1 influenza virus in 2009, which contained the genes from human, avian, and swine influenza viruses, is a representative example of emerging viruses that are recombined and adapted in pigs before transmission to humans (Smith et al., 2009). In our study, the isolate had the same D at positions 225 of the receptor-binding property of HA protein as the 1918 human strain, suggesting that the virus may have potential to infect humans.

The antigenic evolution of the influenza virus via genetic processes of antigenic drift and shift increases antigen variably and leads to epidemics and pandemics. Antigenic drift usually occurs in the antibody-binding sites in the HA protein, the NA protein, or both. It is responsible for the selection pressure to evade host immunity. Lacking immunity in the newly drifted virus will result in a more severe, early-onset influenza epidemic, and increased mortality (Zambon, 2001; Carrat and Flahault, 2007). In our results, we found some changes in the antigenic

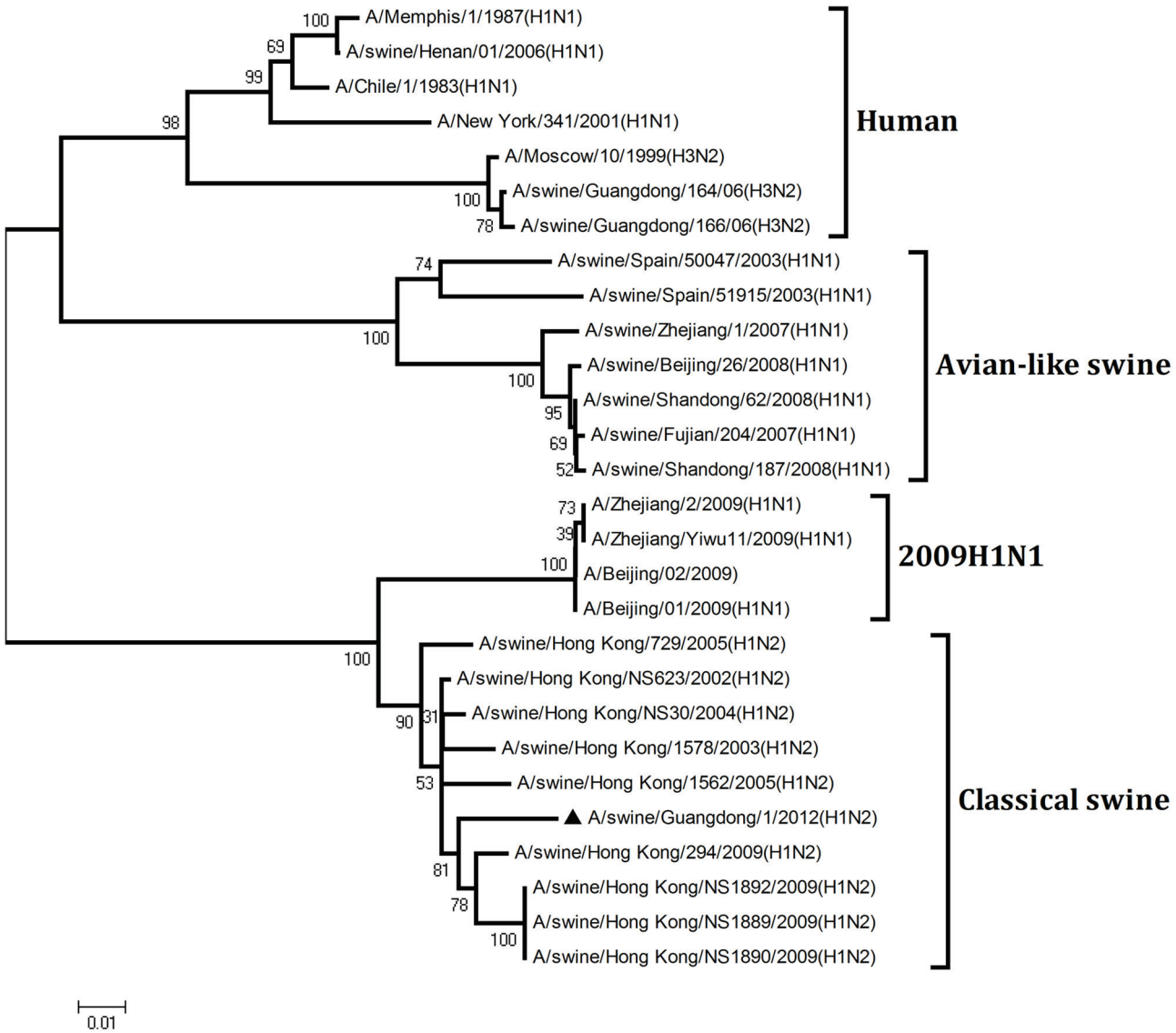


FIGURE 8 | Phylogenetic trees of the PB2, PB1, PA, HA, NP, NA, M, and NS genes of influenza A viruses. The virus marked with "▲" is the swine H1N2 virus [A/swine/Guangdong/1/2012 (H1N2)] isolated and sequenced in this study. The unrooted phylogenetic trees were generated by the distance-based neighbor-joining method using MEGA 4.0 software based on the following sequences: PB2 gene (1), nt 1–2280; PB1 gene (2), nt 1–2274; PA gene (3), nt 1–2151; HA gene (4), nucleotides (nt) 33–1733; NP gene (5), nt 46–1542; NA gene (6), nt 20–1429; M1 (7) gene, nt 11–982; NS1 gene (8), nt 1–660. Bootstrap value was calculated on 1000 replicates of the alignment.

sites of the A/swine/Guangdong/2012 virus: G158E at site Sa; S188N, A189I, and D204E at site Sb; S140P, H141Y, N145R, I169L, S206T, K224R, and Q240G at site Ca; and S74F and N77S at site Cb (H3 numbering). More studies should be done to determine whether these changes in antigenic sites would prompt the reassorted swine virus to infect swine and other hosts.

The amino acid at 627 of PB2 is a key factor for host range, and all human influenza viruses (H1N1, H2N2, and H3N2) have K at this position, whereas the majority of avian influenza viruses have E (Qi et al., 2009). 271A with the 590/591 SR polymorphism in PB2 protein helps pH1N1 and triple-reassortant swine influenza viruses overcome host restriction and efficient replication and adaptation in mammals (Liu et al., 2012). When avian-like signature 627E remains stable rather than changing to the mammalian-like signature 627K, a compensatory D701N substitution increased the polymerase activity and enhanced virulence in mice and enhanced transmission between guinea pigs (Li et al., 2006; Zhou et al., 2013). In our study, the A/swine/Guangdong/1/2012 contained 271A, 590S, 591R, 627E, and 701D in the PB2 protein. The 271A, 590S, and 591R may help our isolate overcome host restriction and efficient

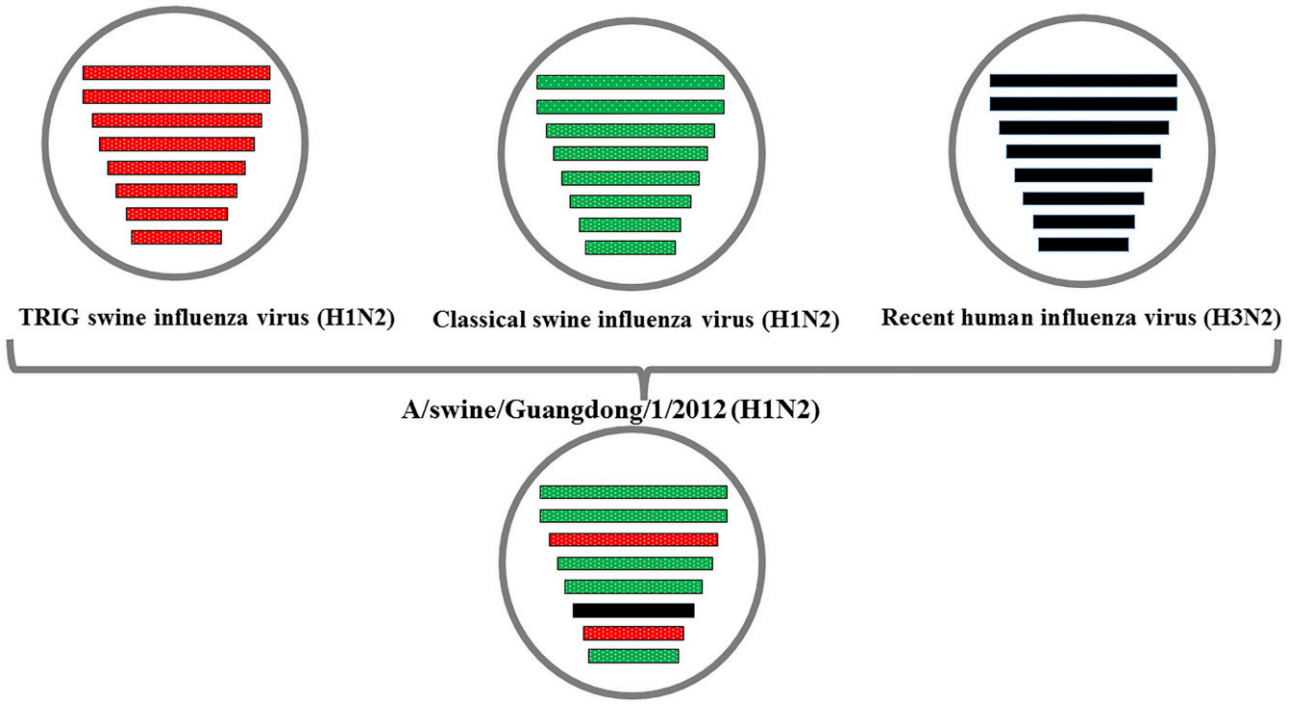


FIGURE 9 | Putative genomic compositions of the novel swine influenza (H1N2) virus (A/swine/Guangdong/1/2012). The eight genes of these viruses were represented by horizontal bars. From top to bottom, PB2, PB1, PA, HA, NP, NA, M, and NS are indicated. Each different color represents a distinct origin.

replication and adaptation in mammals. PB1-F2 is encoded in an alternative reading frame of the PB1 gene and a small protein which is transported to the mitochondria and nucleus. PB1-F2 could induce apoptosis in the host cell and increase virulence and the risk of secondary infections (Chen et al., 2001; McAuley et al., 2007). PB1-F2 protein has variable sizes with truncations either at the C- or N-terminal ends (Vasin et al., 2014). Studies demonstrated that H5N1influenza A virus containing a PB1-F2 was more virulent for BALB/c mice than a closely related H5N1 containing intact PB1-F2 (Kamal et al., 2015). However, a truncated PB1-F2 did not affect the pathogenesis of H1N1 seasonal influenza virus (Meunier and von Messling, 2012). In our study, the isolate and its potential donor virus both contained a truncated PB1-F2 protein (57aa), the role of truncated PB1-F2 of them should be studied in the future. PA-X protein is expressed from a second open reading frame of the PA gene. Studies demonstrated that PA-X decreased the pathogenicity of pandemic 1918 H1N1 virus, 2009 pandemic H1N1 (pH1N1), and highly pathogenic avian influenza H5N1 viruses in mice by modulating the host response (Jagger et al., 2012; Koçer et al., 2013; Gao et al., 2015; Hu et al., 2015). However, Gao et al. demonstrated that PA-X protein in H9N2 virus was a pro-virulence factor in facilitating viral pathogenicity (Gao et al., 2015). In our study, the novel H1N2 and its potential donor virus both possessed a full-length PA-X. The pro- or anti-virulence role of PA-X of them should be studied in the future.

The prevention and control of influenza virus mainly relies on antiviral drugs and vaccines. Amantadine, an adamantane derivative, is an antiviral compound effective against influenza virus. Despite certain side-effects and a rapid induction of resistant strains, amantadine is licensed for the prophylaxis and therapy of influenza in various countries. It inhibits the function of the influenza virus M2 proton channel and single amino acid substitutions at positions L26F, V27A(T), A30T, S31N, and G34E of the M2 protein to confer resistance against it. Single mutant with S31N or double mutants with the S31N and either of the L26F, V27A, or V27T substitutions both confers amantadine resistance (Abed et al., 2005; Krumbholz et al., 2009), However, the significance of V27I exchanges need further study. On the other hand, oseltamivir is used as an NA inhibitor in the treatment of infecting influenza H5N1viruses. Moreover, if the substitutions E119E, H274Y, R292K, and N284S in the NA protein happen, the influenza virus may not be sensitive to the NA inhibitors. In our study, the isolate was still sensitive to NA inhibitors.

Although swine influenza is widespread and is endemic throughout the world, isolating swine influenza viruses is relatively difficult and is dependent on the time of sampling. And the continual co-circulation of antigenically diverse swine influenza virus is a challenge to the production of efficacious and protective vaccines. Antigen-specific antibodies induced by current vaccines provide limited cross protection to heterologous challenges. Moreover, many studies have demonstrated that the

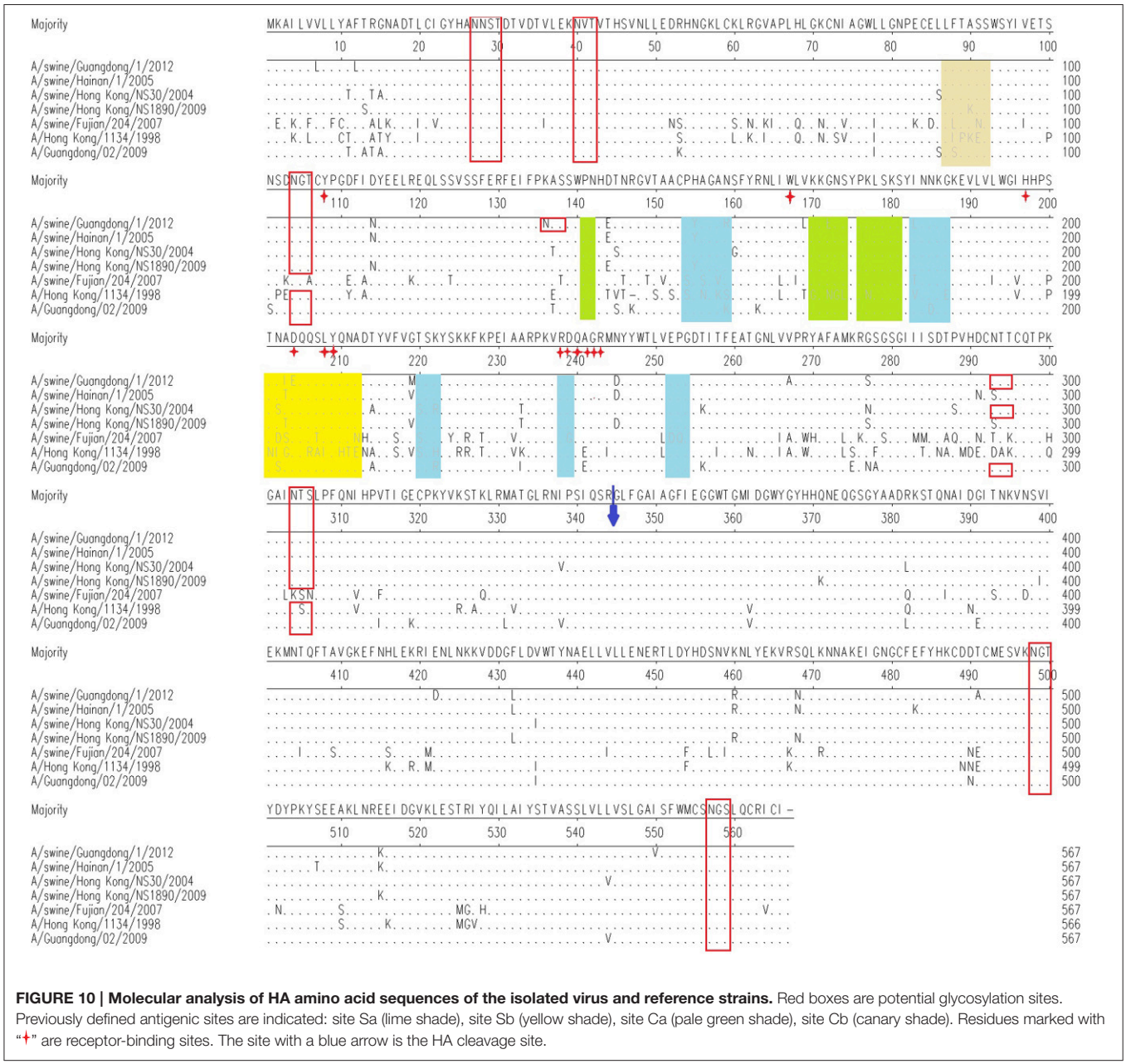


FIGURE 10 | Molecular analysis of HA amino acid sequences of the isolated virus and reference strains. Red boxes are potential glycosylation sites. Previously defined antigenic sites are indicated: site Sa (lime shade), site Sb (yellow shade), site Ca (pale green shade), site Cb (canary shade). Residues marked with "+" are receptor-binding sites. The site with a blue arrow is the HA cleavage site.

vaccine is correlated with an increased risk of influenza-like-illness in swine. Thus, this is why the vaccine is rarely used in swine populations in China. Therefore, it is important to develop new vaccines with high efficacy and safety to protect the swine from influenza viruses in the future.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: PJ. Performed the experiments: YS, XW. Analyzed the data: PJ, YS, XW. Contributed reagents/materials/analysis tools: PJ, XW, YS, NW, GO, NQ, JC. Wrote the paper: YS, YQ, PJ. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (No.31172343), the Science and Technology Projects of Guangdong Province (No.2012B020306003), the Science and Technology Projects of Guangzhou City (No. 201300000037) and the Earmarked Fund for Modern Agro-Industry Technology Research System (CARS-42-G09), Science and Technology Project of the General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China (No. 2010IK023), Science and Technology Project of the Guangdong Entry-Exit Inspection and Quarantine Bureau (No. 2011GDK46).

REFERENCES

Abed, Y., Goyette, N., and Boivin, G. (2005). Generation and characterization of recombinant influenza A (H1N1) viruses harboring amantadine resistance mutations. *Antimicrob. Agents Chemother.* 49, 556–559. doi: 10.1128/AAC.49.2.556-559.2005

Brown, I. H. (2001). "The pig as an intermediate host for influenza A viruses between birds and humans," in *Options for the Control of Influenza Iv* (Crete) 1219, 173–178.

Carrat, F., and Flahault, A. (2007). Influenza vaccine: the challenge of antigenic drift. *Vaccine* 25, 6852–6862. doi: 10.1016/j.vaccine.2007.07.027

Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., et al. (2001). A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* 7, 1306–1312. doi: 10.1038/nm1201-1306

Dawood, F. S., Jain, S., Finelli, L., Shaw, M. W., Lindstrom, S., Garten, R. J., et al. (2009). Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* 360, 2605–2615. doi: 10.1056/NEJMoa0903810

Gao, H., Sun, Y., Hu, J., Qi, L., Wang, J., Xiong, X., et al. (2015). The contribution of PA-X to the virulence of pandemic 2009 H1N1 and highly pathogenic H5N1 avian influenza viruses. *Sci. Rep.* 5:8262. doi: 10.1038/srep08262

Hu, J., Mo, Y., Wang, X., Gu, M., Hu, Z., Zhong, L., et al. (2015). PA-X decreases the pathogenicity of highly pathogenic H5N1 influenza A virus in avian species by inhibiting virus replication and host response. *J. Virol.* 89, 4126–4142. doi: 10.1128/JVI.02132-14

Jagger, B. W., Wise, H. M., Kash, J. C., Walters, K. A., Wills, N. M., Xiao, Y. L., et al. (2012). An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337, 199–204. doi: 10.1126/science.1222213

Kamal, R. P., Kumar, A., Davis, C. T., Tzeng, W. P., Nguyen, T., Donis, R. O., et al. (2015). Emergence of highly pathogenic avian influenza A(H5N1) virus PB1-F2 variants and their virulence in BALB/c mice. *J. Virol.* 89, 5835–5846. doi: 10.1128/JVI.03137-14

Koçer, Z. A., Jones, J. C., and Webster, R. G. (2013). Emergence of influenza viruses and crossing the species barrier. *Microbiol. Spectr.* doi: 10.1128/microbiolspec.OH-0010-2012

Kong, W., Ye, J., Guan, S., Liu, J., and Pu, J. (2014). Epidemic status of Swine influenza virus in china. *Indian J. Microbiol.* 54, 3–11. doi: 10.1007/s12088-013-0419-7

Kothalawala, H., Toussaint, M. J., and Gruys, E. (2006). An overview of swine influenza. *Vet. Q.* 28, 46–53. doi: 10.1080/01652176.2006.9695207

Krumbholz, A., Schmidtke, M., Bergmann, S., Motzke, S., Bauer, K., Stech, J., et al. (2009). High prevalence of amantadine resistance among circulating European porcine influenza A viruses. *J. Gen. Virol.* 90, 900–908. doi: 10.1099/vir.2008.007260-0

Kuiken, T., Holmes, E. C., McCauley, J., Rimmelzwaan, G. F., Williams, C. S., and Grenfell, B. T. (2006). Host species barriers to influenza virus infections. *Science* 312, 394–397. doi: 10.1126/science.1122818

Li, Z., Jiang, Y., Jiao, P., Wang, A., Zhao, F., Tian, G., et al. (2006). The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. *J. Virol.* 80, 11115–11123. doi: 10.1128/JVI.00993-06

Liu, Q., Qiao, C., Marjuki, H., Bawa, B., Ma, J., Guillossou, S., et al. (2012). Combination of PB2 271A and SR polymorphism at positions 590/591 is critical for viral replication and virulence of swine influenza virus in cultured cells and *in vivo*. *J. Virol.* 86, 1233–1237. doi: 10.1128/JVI.05699-11

McAuley, J. L., Hornung, F., Boyd, K. L., Smith, A. M., McKeon, R., Bennink, J., et al. (2007). Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* 2, 240–249. doi: 10.1016/j.chom.2007.09.001

Meunier, I., and von Messling, V. (2012). PB1-F2 modulates early host responses but does not affect the pathogenesis of H1N1 seasonal influenza virus. *J. Virol.* 86, 4271–4278. doi: 10.1128/JVI.07243-11

Nichol, S. T., Arikawa, J., and Kawaoka, Y. (2000). Emerging viral diseases. *Proc. Natl. Acad. Sci. U.S.A.* 97, 12411–12412. doi: 10.1073/pnas.210382297

Qi, X., and Lu, C. P. (2006). Genetic characterization of novel reassortant H1N2 influenza A viruses isolated from pigs in southeastern China. *Arch. Virol.* 151, 2289–2299. doi: 10.1007/s00705-006-0796-x

Qi, X., Pang, B., and Lu, C. P. (2009). Genetic characterization of H1N1 swine influenza A viruses isolated in eastern China. *Virus Genes* 39, 193–199. doi: 10.1007/s11262-009-0375-9

Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459, 1122–1125. doi: 10.1038/nature08182

Song, Y., Cui, J., Song, H., Ye, J., Zhao, Z., Wu, S., et al. (2015). New reassortant H5N8 highly pathogenic avian influenza virus from waterfowl in Southern China. *Front. Microbiol.* 6:1170. doi: 10.3389/fmicb.2015.01170

Sugimura, T., Yonemochi, H., Ogawa, T., Tanaka, Y., and Kumagai, T. (1980). Isolation of a recombinant influenza virus (Hsw 1 N2) from swine in Japan. *Arch. Virol.* 66, 271–274. doi: 10.1007/BF01314741

Vasin, A. V., Temkina, O. A., Egorov, V. V., Klotchenko, S. A., Plotnikova, M. A., and Kiselev, O. I. (2014). Molecular mechanisms enhancing the proteome of influenza A viruses: an overview of recently discovered proteins. *Virus Res.* 185, 53–63. doi: 10.1016/j.virusres.2014.03.015

Vijaykrishna, D., Smith, G. J., Pybus, O. G., Zhu, H., Bhatt, S., Poon, L. L., et al. (2011). Long-term evolution and transmission dynamics of swine influenza A virus. *Nature* 473, 519–522. doi: 10.1038/nature10004

Yu, H., Zhang, P. C., Zhou, Y. J., Li, G. X., Pan, J., Yan, L. P., et al. (2009). Isolation and genetic characterization of avian-like H1N1 and novel ressortant H1N2 influenza viruses from pigs in China. *Biochem. Biophys. Res. Commun.* 386, 278–283. doi: 10.1016/j.bbrc.2009.05.056

Zambon, M. C. (2001). The pathogenesis of influenza in humans. *Rev. Med. Virol.* 11, 227–241. doi: 10.1002/rmv.319

Zhou, B., Pearce, M. B., Li, Y., Wang, J., Mason, R. J., Tumpey, T. M., et al. (2013). Asparagine substitution at PB2 residue 701 enhances the replication, pathogenicity, and transmission of the 2009 pandemic H1N1 influenza A virus. *PLoS ONE* 8:e67616. doi: 10.1371/journal.pone.0067616

Zhu, H., Webby, R., Lam, T. T., Smith, D. K., Peiris, J. S., and Guan, Y. (2013). History of Swine influenza viruses in Asia. *Curr. Top. Microbiol. Immunol.* 370, 57–68. doi: 10.1007/82_2011_179

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Song, Wu, Wang, Ouyang, Qu, Cui, Qi, Liao and Jiao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 April 2016
doi: 10.3389/fmicb.2016.00520

Continuing Reassortant of H5N6 Subtype Highly Pathogenic Avian Influenza Virus in Guangdong

*Runyu Yuan[1,2,3], Zheng Wang[4], Yinfeng Kang[3], Jie Wu[1,2], Lirong Zou[1,2], Lijun Liang[1,2], Yingchao Song[1,2], Xin Zhang[1,2], Hanzhong Ni[1,2], Jinyan Lin[1,2] and Changwen Ke[1,2]**

[1] Key Laboratory for Repository and Application of Pathogenic Microbiology, Research Center for Pathogens Detection Technology of Emerging Infectious Diseases, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China, [2] WHO Collaborating Centre for Surveillance, Research and Training of Emerging Infectious Disease, Guangzhou, China, [3] Key Laboratory of Zoonosis Prevention and Control of Guangdong, College of Veterinary Medicine, South China Agricultural University, Guangzhou, China, [4] School of Public Health, Sun Yat-Sen University, Guangzhou, China

First identified in May 2014 in China's Sichuan Province, initial cases of H5N6 avian influenza virus (AIV) infection in humans raised great concerns about the virus's prevalence, origin, and development. To evaluate both AIV contamination in live poultry markets (LPMs) and the risk of AIV infection in humans, we have conducted surveillance of LPMs in Guangdong Province since 2013 as part of environmental sampling programs. With environmental samples associated with these LPMs, we performed genetic and phylogenetic analyses of 10 H5N6 AIVs isolated from different cities of Guangdong Province from different years. Results revealed that the H5N6 viruses were reassortants with hemagglutinin (HA) genes derived from clade 2.3.4.4 of H5-subtype AIV, yet neuraminidase (NA) genes derived from H6N6 AIV. Unlike the other seven H5N6 viruses isolated in first 7 months of 2014, all of which shared remarkable sequence similarity with the H5N1 AIV in all internal genes, the PB2 genes of GZ693, GZ670, and ZS558 more closely related to H6N6 AIV and the PB1 gene of GZ693 to the H3-subtype AIV. Phylogenetic analyses revealed that the environmental H5N6 AIV related closely to human H5N6 AIVs isolated in Guangdong. These results thus suggest that continued reassortment has enabled the emergence of a novel H5N6 virus in Guangdong, as well as highlight the potential risk of highly pathogenic H5N6 AIVs in the province.

Keywords: reassortant, highly pathogenicity, avian influenza virus, H5N6, live poultry market

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University Graduate School, Japan

Reviewed by:
Julie McAuley,
University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Australia
Ram P. Kamal,
Battelle, USA
Karoline Bragstad,
The Norwegian Institute of Public Health, Norway

****Correspondence:***
Changwen Ke
kecw1965@aliyun.com

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *04 February 2016*
Accepted: *29 March 2016*
Published: *13 April 2016*

Citation:
Yuan R, Wang Z, Kang Y, Wu J, Zou L, Liang L, Song Y, Zhang X, Ni H, Lin J and Ke C (2016) Continuing Reassortant of H5N6 Subtype Highly Pathogenic Avian Influenza Virus in Guangdong. Front. Microbiol. 7:520. doi: 10.3389/fmicb.2016.00520

INTRODUCTION

Depending on their pathotypes, avian influenza viruses (AIVs) have inherently different pathogeneses in the infection and distribution of lesions. According to pathotype, AIVs can divide into non-pathogenic AIVs (NPAIV), low pathogenic AIVs (LPAIV), and highly pathogenic AIVs (HPAIV). In birds, NPAIVs (e.g., H6N6) usually present no clinical symptoms; by contrast, LPAIVs (e.g., H9N2) can cause mild respiratory or gastrointestinal infection, and HPAIVs (e.g., H5N1 and H7N2) can induce systemic, multi-organ infection, as well as high morbidity, and mortality (Swayne and Halvorson, 2003; Yuan et al., 2014).

A continual threat to animal and human health, HPAIVs have caused infections and deaths in not only countless birds, but also many humans. In 1997, the H5N1 HPAIV, the

internal genes of which derived from the H6N1 NPAIV, infected 18 people in Hong Kong, six of whom died from the infection (Claas et al., 1998; Subbarao et al., 1998). Furthermore, since 2003, the H5N1 HPAIV has caused outbreaks both in birds and humans in more than 60 countries, including China (Yuan et al., 2014; WHO, 2015a). Recently, H5-subtype HPAIV s—that is, variants of different NA subtypes—have also caused outbreaks in poultry in China (i.e., subtypes H5N1, H5N2, H5N5, H5N6, and H5N8), as well as in South Korea (i.e., subtype H5N8), Japan (i.e., subtype H5N8), Laos (i.e., subtype H5N8), and Vietnam (i.e., subtypes H5N1 and H5N6; WHO, 2014d; OIE, 2015). In March 2014, an outbreak of H5N6 HPAIV in poultry was reported in Laos and, that April, in Vietnam (Wong et al., 2015). Genetic studies have shown that the H5N6 virus has exchanged genes from the H5N1 and H6N6 AIVs that circulate widely in ducks (Shen et al., 2015). Although little is known about the potential of these novel viruses to infect humans, a few isolated cases have been detected. On May 6, 2014, one such case of H5N6 infection in China's Sichuan Province was fatal (CDC China, 2014; WHO, 2014c), and later that year, another severe case of infection occurred in Guangdong Province in December (WHO, 2014a). As of February 2016, nine cases of H5N6 AIVs infection in humans have been confirmed in China, six of them in Guangdong Province (WHO, 2015b,c, 2016a,b,c).

Since 2013, several surveillance systems for pandemic preparedness have been established in China, including those at live poultry markets (LPM) and sentinel hospitals. These surveillance systems have played a vital role in the early detection of warning signs of AIV infection in humans. During our study's surveillance period, we isolated 10 H5N6 AIVs in environmental samples from LPMs in Guangdong Province, and to better understand their genetic diversity and evolution, we analyzed their related epidemiological and sequence data.

MATERIALS AND METHODS

Ethics Statement

This research was reviewed and approved by the South China Agricultural University Experimental Animal Welfare Ethics Committee (permit no. 2014-11).

Sample Collection

Beginning on April 16, 2013, in order to better monitor LPMs for AIV contamination and assess the risk of AIV infection in humans, environmental sampling programs were implemented in Guangdong Province. Environmental samples were taken from poultry excrement, epilator swabs, and sewage swabs—the latter two from drains in meat preparation areas or around cages—whereas chopping swab samples were gathered randomly from butcher boards or knives at LPMs each week.

Virus Isolation

Samples were first tested for influenza A by using real-time polymerase chain reactions (qPCR) in the laboratories of the district's Centers for Disease Control and Prevention (CDC). Positive influenza A samples were probed to detect subtypes H5, H7, and H9 by using qPCR in local CDC laboratories, and results were later verified by Guangdong's CDC. H5-positive samples were further analyzed by using qPCR to detect the presence of the N6 gene. All qPCR-detected primers and probes were provided by the Chinese CDC. Samples positive with H5N6 subtypes were purified and propagated in 10-d embryonated chicken eggs free of specific pathogens and stored at −70°C until used. Subtypes of the viruses were further identified by hemagglutination (HA) inhibition assay. All experiments were carried out in animal biosafety level 3 facilities.

Genomic Sequencing

Viral RNA was first extracted from allantoic fluid by using an RNA extraction kit (QIAamp Viral RNA Mini Kit, Qiagen, Hilden, Germany). Reverse transcription and polymerase chain reaction (PCR) amplification of all eight gene segments used pre-amplification reagents (PathAmp^{TM} FluA, Life Technologies, Guilford, Connecticut, USA). PCR products were purified and quantified with a purification kit (AmpureXP, Beckman Coulter, Porterville, CA, USA) according to the manufacturer's instructions. The full genomes of the viruses were sequenced with a sequencing kit (Ion PGM Sequencing 200 Kit version 2, Life Technologies), specifically with the kit's Ion 316 Chip V2 and according to the manufacturer's instructions.

Sequence Analysis

To align and analyze the sequences, multiple sequences of the representative AIVs were downloaded from GenBank databases (Li et al., 2010; Yuan et al., 2014, 2016). Full-length gene sequences were implemented and edited with Lasergene 7.1 (DNASTAR, Madison, Wisconsin, USA). A neighbor-joining algorithm and maximum-likelihood trees model were estimated for all eight genes—namely, HA, NA, PB2, PB1, PA, NP, M, and NS—by using genetic analysis software [Molecular Evolutionary Genetics Analysis (MEGA) version 6.06] with 1000 bootstrap trials. Branches with bootstrap values exceeding 50% were grouped together in the trees.

Nucleotide sequences obtained in our study, all listed by their accession numbers, are currently available from GenBank (**Table 1**).

RESULTS

Prevalence of the H5-Subtype AIV in LPMs

From April 2013 to December 2015, a total of 32,452 fecal and swabs were collected from LPMs in 21 cities in Guangdong Province (**Table 2**). Among all of the samples, 6865 (21.2%) were positive for influenza A, 14.6% of which with the H5 subtype. The H5N1 subtype was the most prevalent among the H5 subtypes, followed by H5N6; also observed were H5N2, H5N3, H5N4, H5N5, H5N7, H5N8, and H5N9. During the same period, we selected 10 H5N6 subtypes among the 66 H5N6-positive samples in different cities of Guangdong Province in different years to analyze the evolution of the subtype (**Table 1**).

Phylogenetic Analysis of Surface Genes

The genomes of the 10 H5N6 AIVs isolated from environmental samples were sequenced by using a next-generation sequencer

TABLE 1 | Isolation of H5N6-subtype avian influenza viruses from live poultry markets in Guangdong, 2013–2015.

Virus	Abbreviation	Collection city	Collection date	Accession number (PB2, PB1, PA, HA, NP, NA, M, NS)
A/Environment/Guangdong/QY025/2013(H5N6)	QY025	Qingyuan	2013.04	KT370109, KT370102, KT370095, KT370059, KT370082, KT370076, KT370067, KT370088
A/Environment/Guangdong/JY137/2014(H5N6)	JY137	Jieyang	2014.03	KT370111, KT370105, KT370098, KT370064, KT370084, KT370074, KT370071, KT370092
A/Environment/Guangdong/PY955/2014(H5N6)	PY955	Guangzhou	2014.12	KT370113, KT370106, KT370099, KT370062, KT370079, KT370075, KT370069, KT370091
A/Environment/Guangdong/QY197/2014(H5N6)	QY197	Qingyuan	2014.05	KT370108, KT370101, KT370094, KT370060, KT370081, KT370077, KT370066, KT370087
A/Environment/Guangdong/QY208/2014(H5N6)	QY208	Qingyuan	2014.05	KT370110, KT370103, KT370096, KT370058, KT370083, KT370073, KT370068, KT370089
A/Environment/Guangdong/ZS356/2014(H5N6)	ZS356	Zhongshan	2014.07	KT370107, KT370100, KT370093, KT370061, KT370080, KT370078, KT370065, KT370086
A/Environment/Guangdong/HY243/2015(H5N6)	HY243	Heyuan	2015.02	KT370112, KT370104, KT370097, KT370063, KT370085, KT370072, KT370070, KT370090
A/Environment/Guangdong/GZ670/2015(H5N6)	GZ670	Guangzhou	2015.10	KU852961, KU852958, KU852955, KU852946, KU852952, KU852949, KU852964, KU852965
A//Environment/Guangdong/GZ693/2015(H5N6)	GZ693	Guangzhou	2015.11	KU852959, KU852956, KU852953, KU852944, KU852950, KU852947, KU852962, KU852966
A/Environment/Guangdong/ZS558/2015(H5N6)	ZS558	Zhongshan	2015.05	KU852960, KU852957, KU852954, KU852945, KU852951, KU852948, KU852963, KU852967

(Ion PGM, Life Technologies). The complete genomes of the 10 samples were compared with nucleotide sequences of some viruses in GenBank databases.

Phylogenetic analyses demonstrated the origin and evolution of H5N6 AIVs in China. As results of the phylogenetic analysis of H5 and related viruses show, the HA gene of all 10 viruses clustered into clade 2.3.4.4 (**Figure 1A**) and thus related more closely to the H5N2 HPAIV, A/chicken/Zhejiang/727159/2014(H5N2), which circulates in Zhejiang Province (**Figures 1A**, **2**). In addition, QY025, QY197, QY208, GZ670, and ZS558 shared 98.5–99.9% highest nucleotide similarity with A/chicken/Dongguan/2690/2013(H5N6) (GD-H5N6), JY137, PY955, and ZS356 shared 99.1–99.6% highest nucleotide similarity with A/chicken/Shenzhen/1395/2013(H5N6) (GD-H5N6), and HY243 shared 99.1% highest nucleotide similarity with JX-H5N6. More singularly, GZ693 shared 98.7% highest nucleotide similarity with A/Guangdong/ZQ874/2015(H5N6) (ZQ874), which was found to have recently infected human in Zhaoqing, Guangdong Province.

Phylogenetic analysis of the N6-NA gene indicated that it likely originated in H6N6 AIVs found in domestic ducks in southern China (Huang et al., 2012). According to their geographical location, the HA genes of H6-subtype AIV can be identified as either of Eurasian or North American lineage. As **Figure 1B** shows, those of Eurasian lineage can be divided into two groups: Group 1 (ST192-like) and Group 2 (ST4893-like) (Huang et al., 2012). All 10 viruses were of Group 1 of Eurasian lineage, represented by A/wild duck/Shantou/192/2004 (H6N6) (**Figures 1B**, **2**). QY025, QY197, and QY208 shared 97.2–98.8% nucleotide similarity with A/chicken/Shenzhen/552/2013(H5N6) (GD-H5N6), JY137, ZS356, HY243, and ZS558 shared 96.5–99.6% highest nucleotide similarity with A/chicken/Dongguan/2685/2013(H5N6) (GD-H5N6), PY955 shared 99.6% highest nucleotide similarity with JX-H5N6, and GZ670 and GZ693 shared 96.0–96.9% highest nucleotide similarity with ZQ874 isolated from a patient in Guangdong. In all, these results suggest that the surface genes of the 10 reassortant viruses derived from H5 and H6 AIV subtypes circulating in poultry in China.

Phylogenetic Analysis of Internal Genes

Phylogenetic analyses of internal genes showed that the PB2, PB1, PA, M, and NS genes of all 10 viruses were of Eurasian lineage (**Figures 1C–H**). H5N6 AIVs did not cluster with H5N1 AIVs, but formed an independent lineage (**Figure 1**).

In the PB2 gene, QY025, QY197, QY208, JY137, and ZS356 shared 99.1–99.9% highest nucleotide similarity with GD-H5N6, whereas ZS356, PY955 and HY243 shared more than 99.4% highest nucleotide similarity with JX-H5N6. GZ670, GZ693, and ZS558 originated from H6-subtype AIVs and were ST339-like viruses, represented by A/wild duck/Shantou/339/2000 (H6N2) (**Figure 1C**).

From the PB1 gene, PY955 and HY243 shared 99.0–99.1% highest nucleotide similarity with DG-H5N6, whereas the other seven viruses shared more than 98.8% highest nucleotide similarity with Vietnam-H5N6. Meanwhile, the PB1 gene of

TABLE 2 | Environment surveillance of H5-subtype avian influenza viruses in Guangdong, 2013–2015.

City	2013			2014			2015		
	No. of samples collected	No. of positive samples collected		No. of samples collected	No. of positive samples collected		No. of samples collected	No. of positive samples collected	
		Total	H5		Total	H5		Total	H5
Chaozhou	–[a]	–	–	5	–	–	582	96	0
Dongguan	321	37	23	588	24	13	4305	358	96
Foshan	31	2	0	464	121	12	1837	412	43
Guangzhou	274	2	2	655	39	4	169	90	0
Heyuan	–	–	–	301	13	5	801	65	30
Huizhou	–	–	–	80	56	11	350	141	12
Jiangmen	–	–	–	690	204	17	1292	376	43
Jieyang	–	–	–	311	190	49	683	215	22
Maoming	157	18	3	337	121	18	663	167	29
Meizhou	101	5	0	1050	121	10	1951	187	4
Qingyuan	–	–	–	240	84	13	604	244	27
Shantou	–	–	–	364	136	11	747	63	0
Shanwei	–	–	–	64	14	5	421	35	0
Shaoguan	–	–	–	105	35	1	535	105	7
Shenzhen	196	48	14	384	151	25	621	57	0
Yangjiang	343	69	–	1344	214	3	921	104	10
Yunfu	213	41	7	374	85	5	725	142	5
Zhanjiang	–	–	–	84	32	6	415	45	5
Zhaoqing	405	108	18	110	27	2	1470	381	83
Zhongshan	–	–	–	1021	529	155	1031	324	89
Zhuhai	–	–	–	295	163	13	599	200	7
Total	2041	330	67	8866	2359	378	21,545	4176	554
		(16.2%)[b]	(20.3%)[c]		(26.6%)	(16.0%)		(19.4%)	(13.3%)

[a] *Undetected.*
[b] *Percent positives of total collected samples in a year.*
[c] *Percent H5 positives of total positives samples in a year.*

GZ693 shared highest nucleotide similarity with H3-subtype AIVs.

Regarding the PA gene, ZS356, and HY243 shared 99.7–99.3% highest nucleotide similarity with JX-H5N6, whereas the other eight viruses shared more than 99.3% highest nucleotide similarity with GD-H5N6. As for the NP gene, PY955, ZS356, and HY243 shared more than 99.7% highest nucleotide similarity with JX-H5N6, though the other seven viruses shared 99.4–99.8% highest nucleotide similarity with GD-H5N6, and concerning the M gene, all 10 viruses shared 99.6–100% highest nucleotide similarity with JX-H5N6. Lastly, regarding the NS gene, PY955 and JY137 shared 97.6–99.9% highest nucleotide similarity with JX-H5N6, whereas the other eight viruses shared 98.7–99.4% highest nucleotide similarity with GD-H5N6.

In particular, the 10 environmental viruses shared more than 96.0% high nucleotide similarity with the H5N6 AIVs isolated from patients in Guangdong. Phylogenetic analysis demonstrated that the internal genes of seven AIVs isolated within the first 7 months of 2014 related more closely to H5N1 HPAIVs circulating in poultry in China. By contrast, GZ670, GZ693, and ZS558 isolated in 2015 diverged from previously sequenced H5N6 AIVs and related more closely to H6N2 AIVs in the PB2 gene (**Figure 2**).

Molecular Characterization

The HA gene of all 10 H5N6 AIVs showed the HPAIV amino acid sequence RERRRKR↓G at the cleavage site of HA1 and HA2. Amino acid residues Q226 and G228, according to H3 numbering, occurred in the receptor-binding pocket of HA1, thus indicating that the viruses preferred to bind to the AIV receptor (Ha et al., 2001). Each of the 10 AIVs had six potential N-linked glycosylation sites at HA1 (26 or 27, 39, 181, 209, and 302) and two in HA2 (499 and 558). However, ZS558 revealed A254T mutation in an extra potential glycosylation site, whereas GZ693 exhibited six potential N-linked glycosylation sites in HA1 (i.e., at positions 27, 39, 180, 208, 230, and 301) and two in HA2 (i.e., at positions 498 and 557).

The NA proteins of JY137 and PY955 exhibited 12 amino acid deletion residues (i.e., at positions 59–70) in the neck, which could boost its virulence in mammals (Matsuoka et al., 2009). The key antiviral neuraminidase inhibitor drugs sites of the NA and M genes, such as position H275 of the NA gene (NA of GS/GD

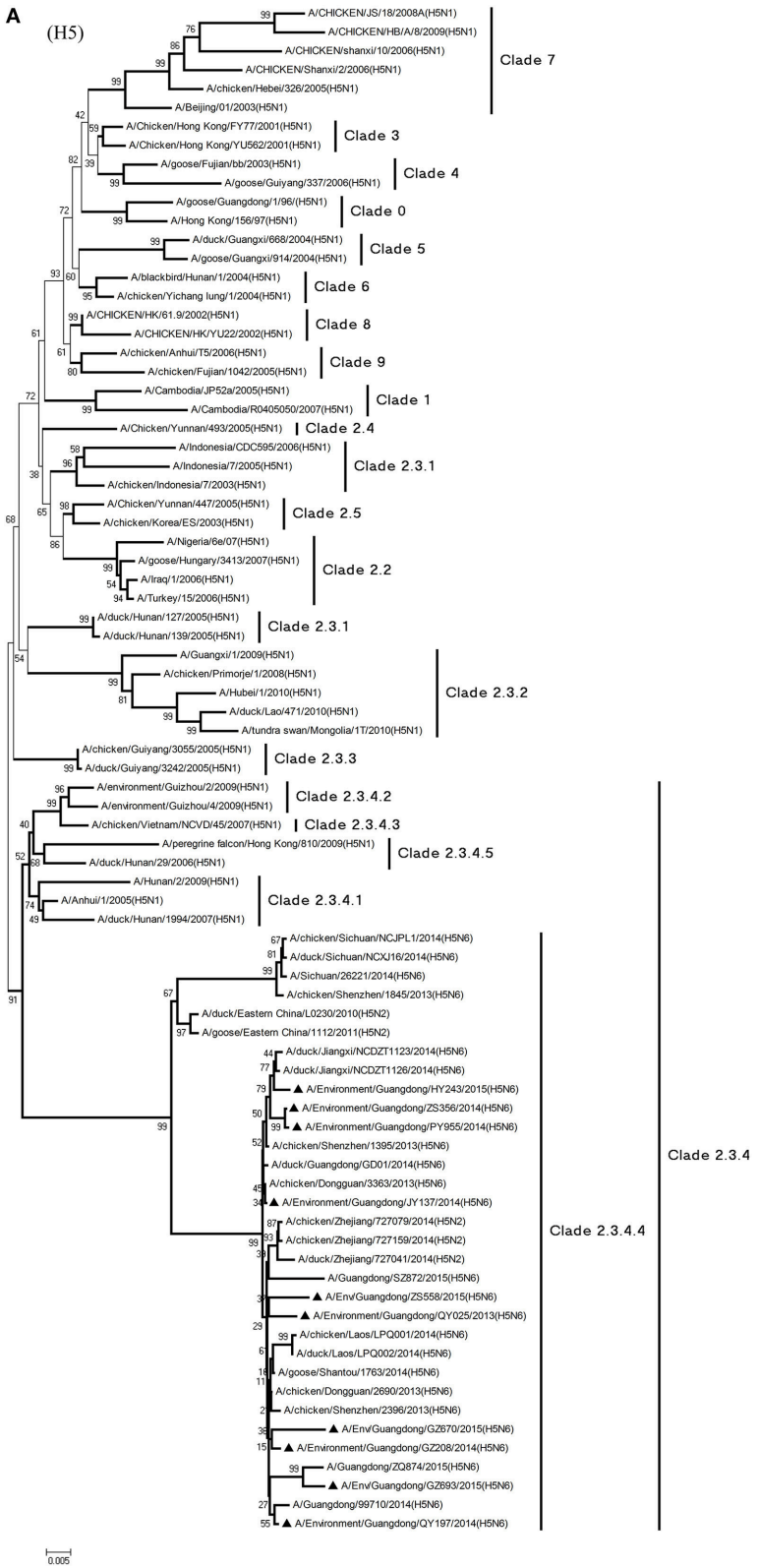


FIGURE 1 | Continued

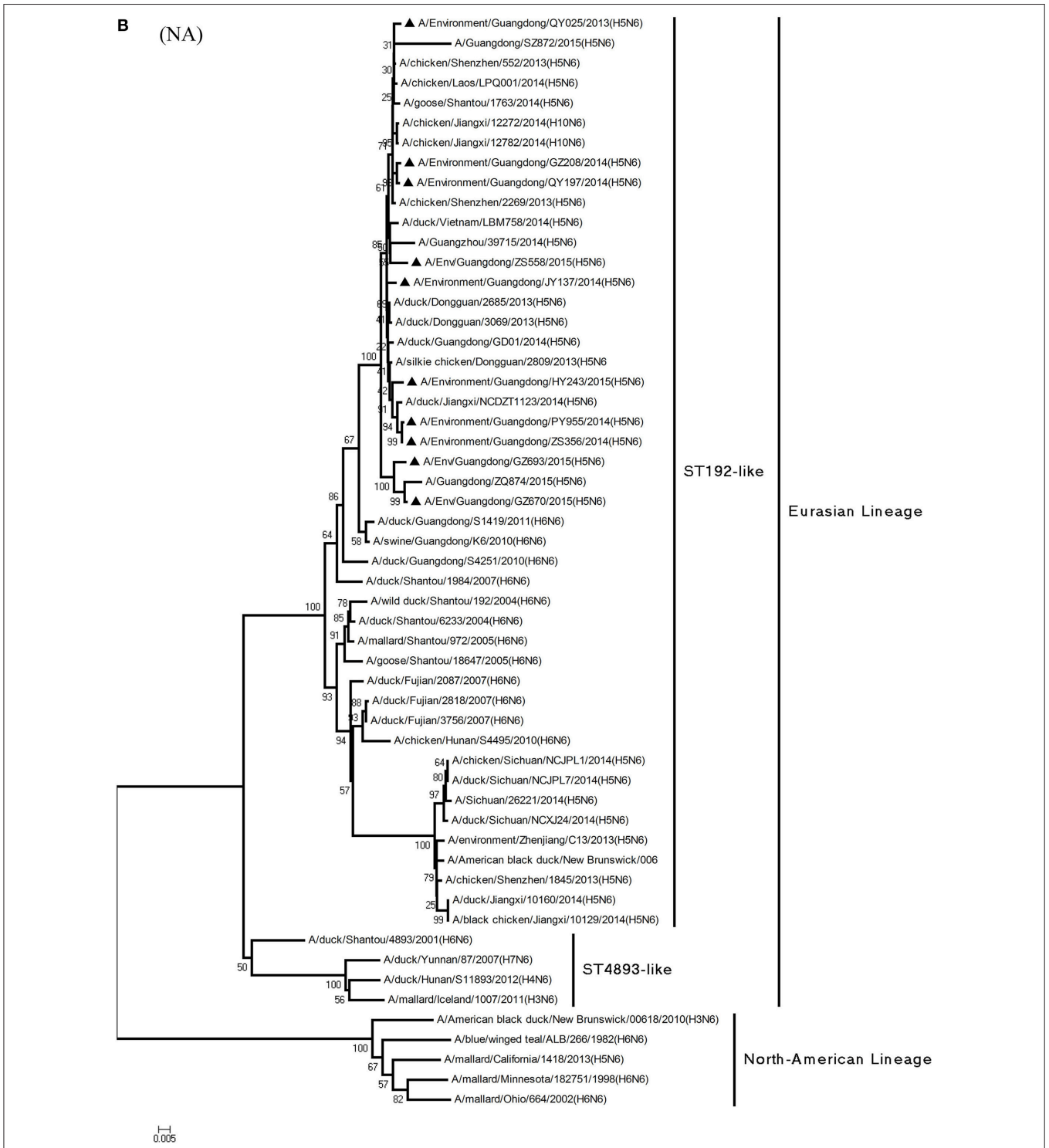


FIGURE 1 | Continued

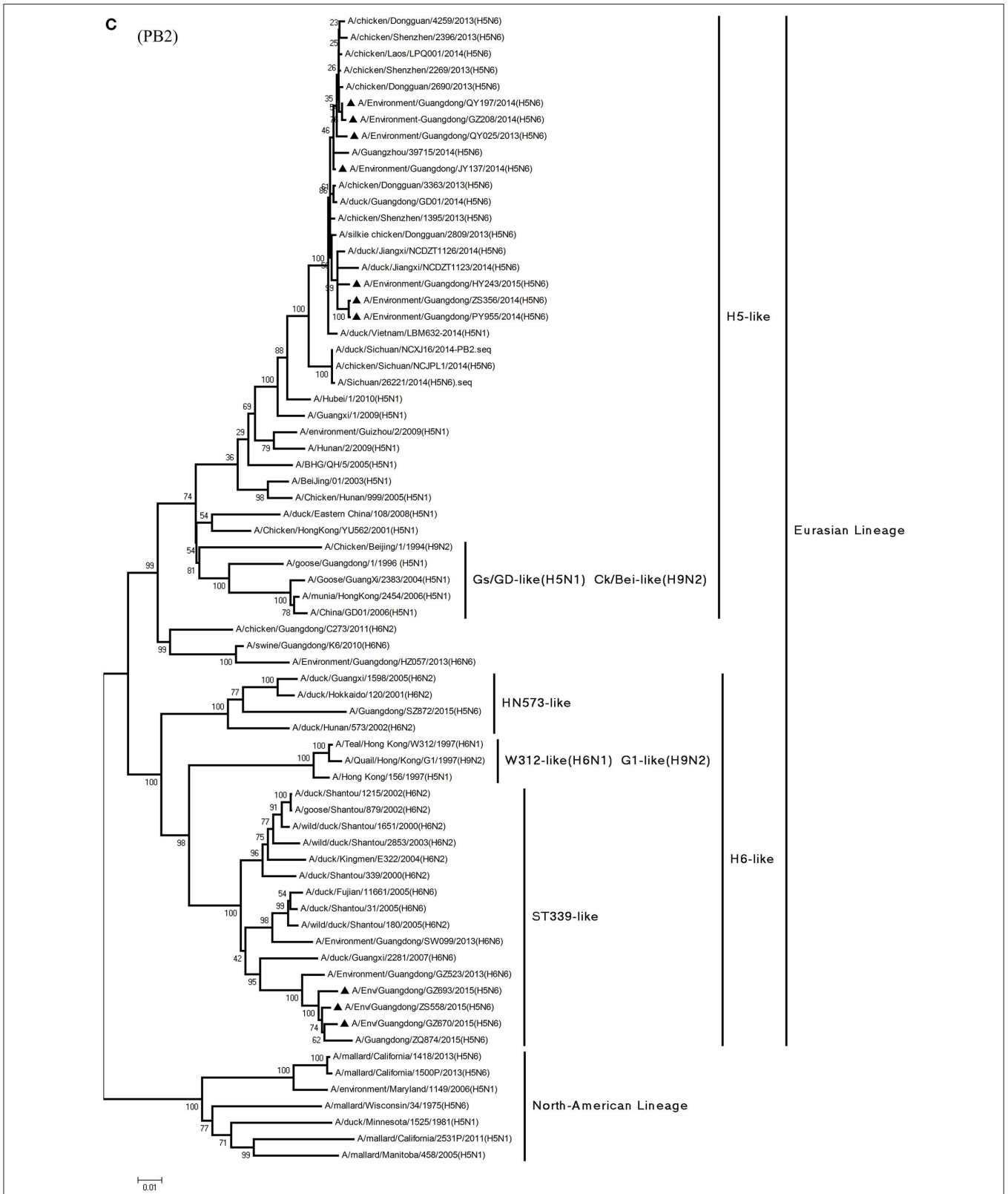


FIGURE 1 | Continued

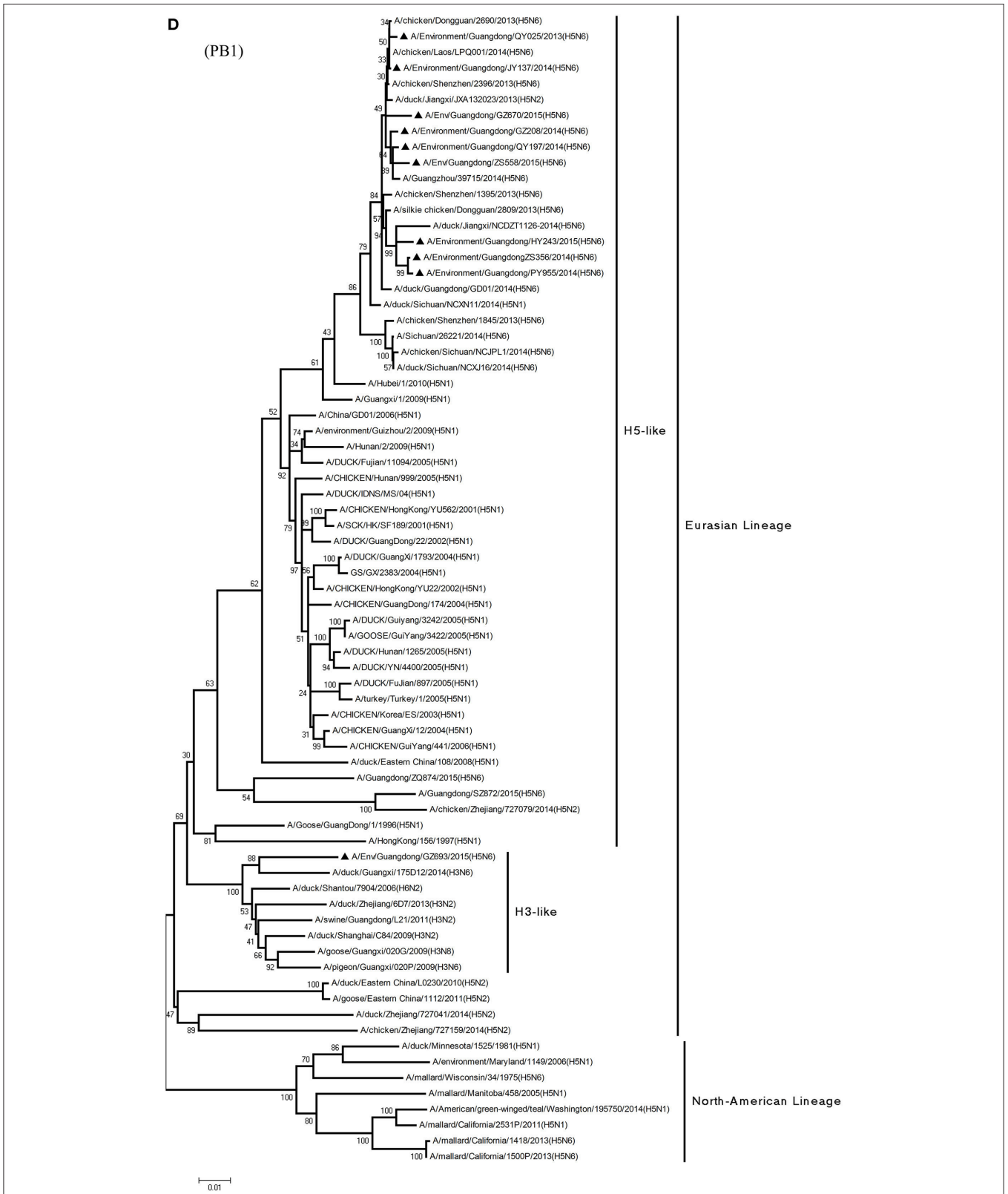


FIGURE 1 | Continued

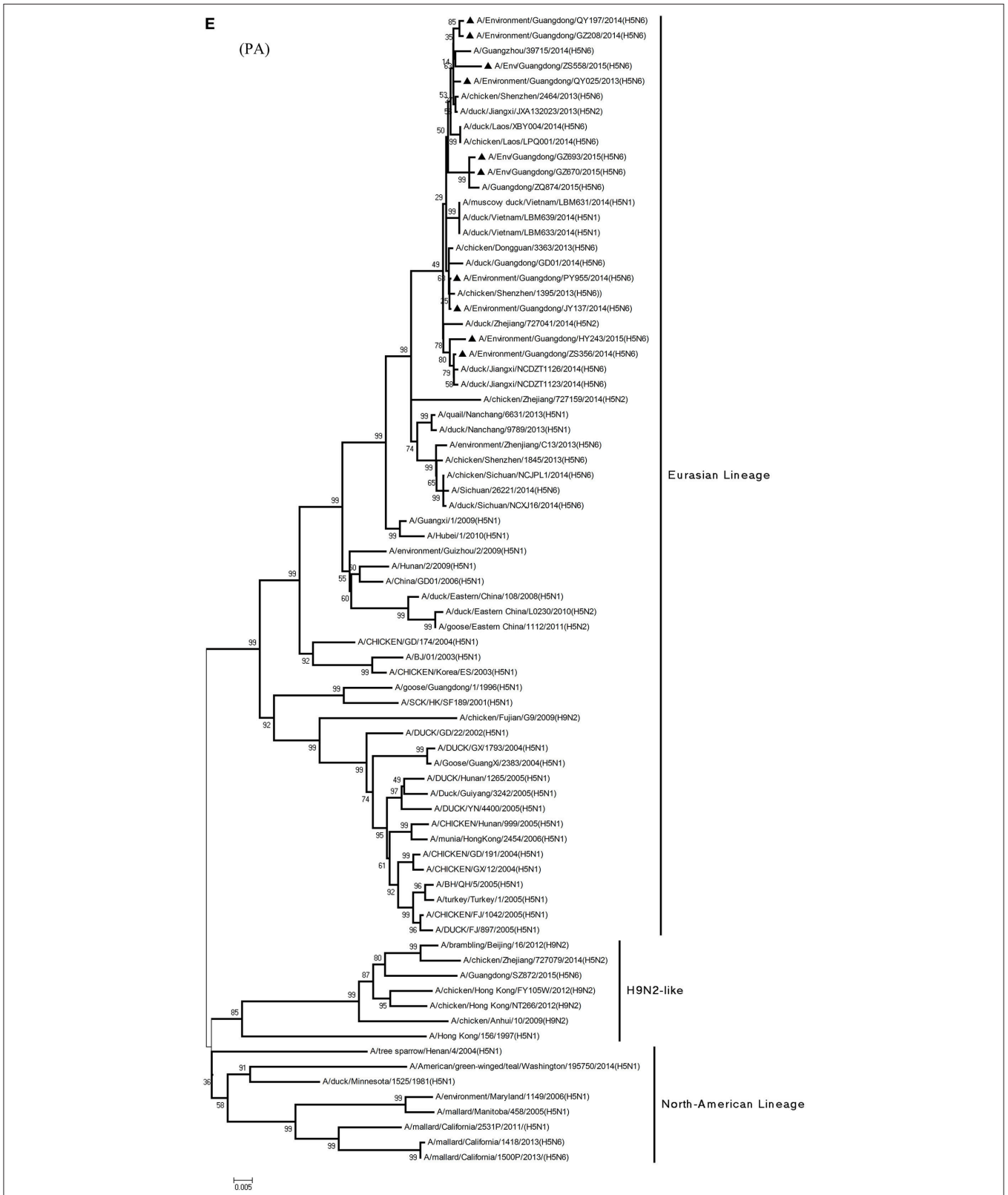


FIGURE 1 | Continued

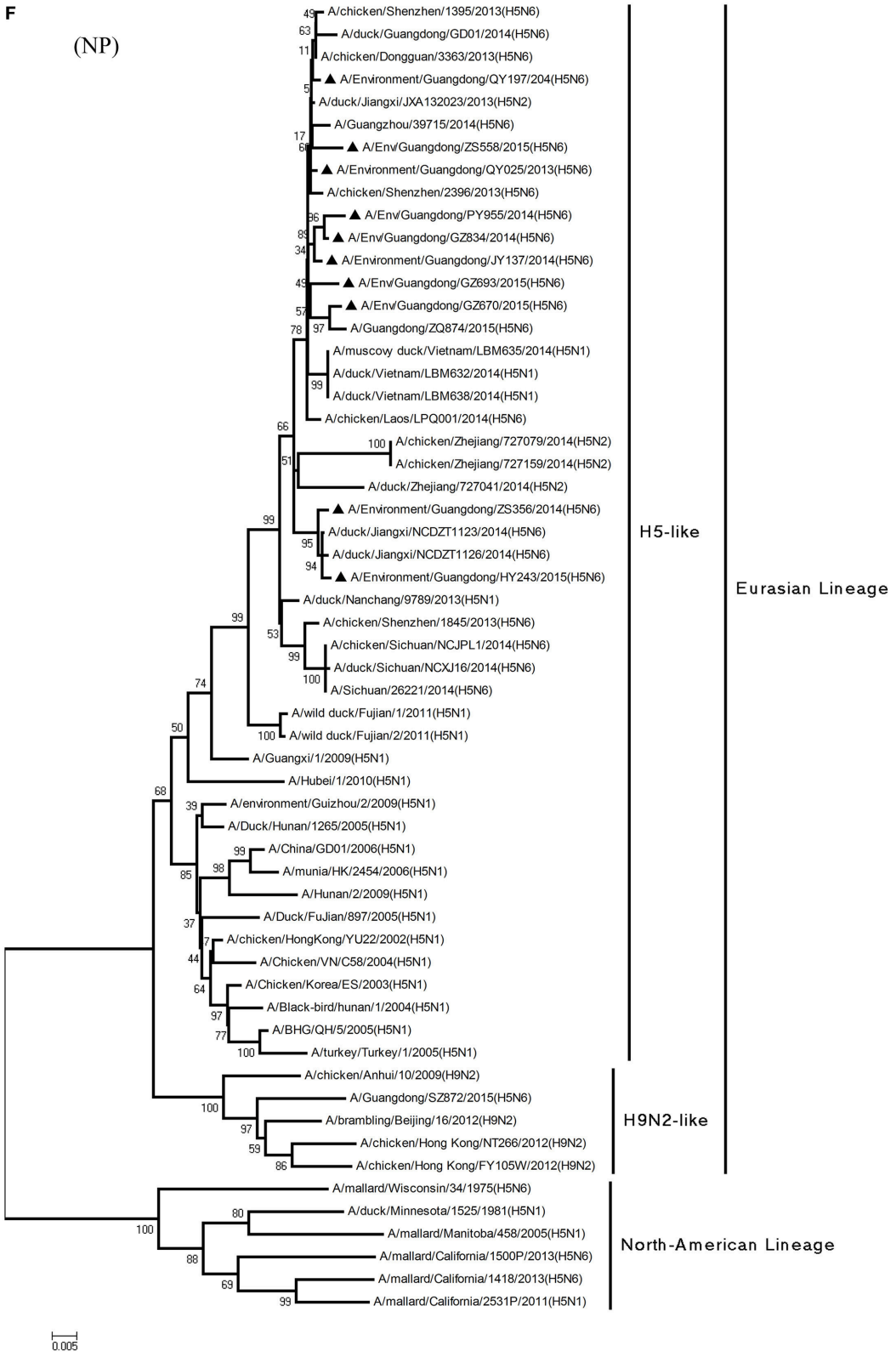


FIGURE 1 | Continued

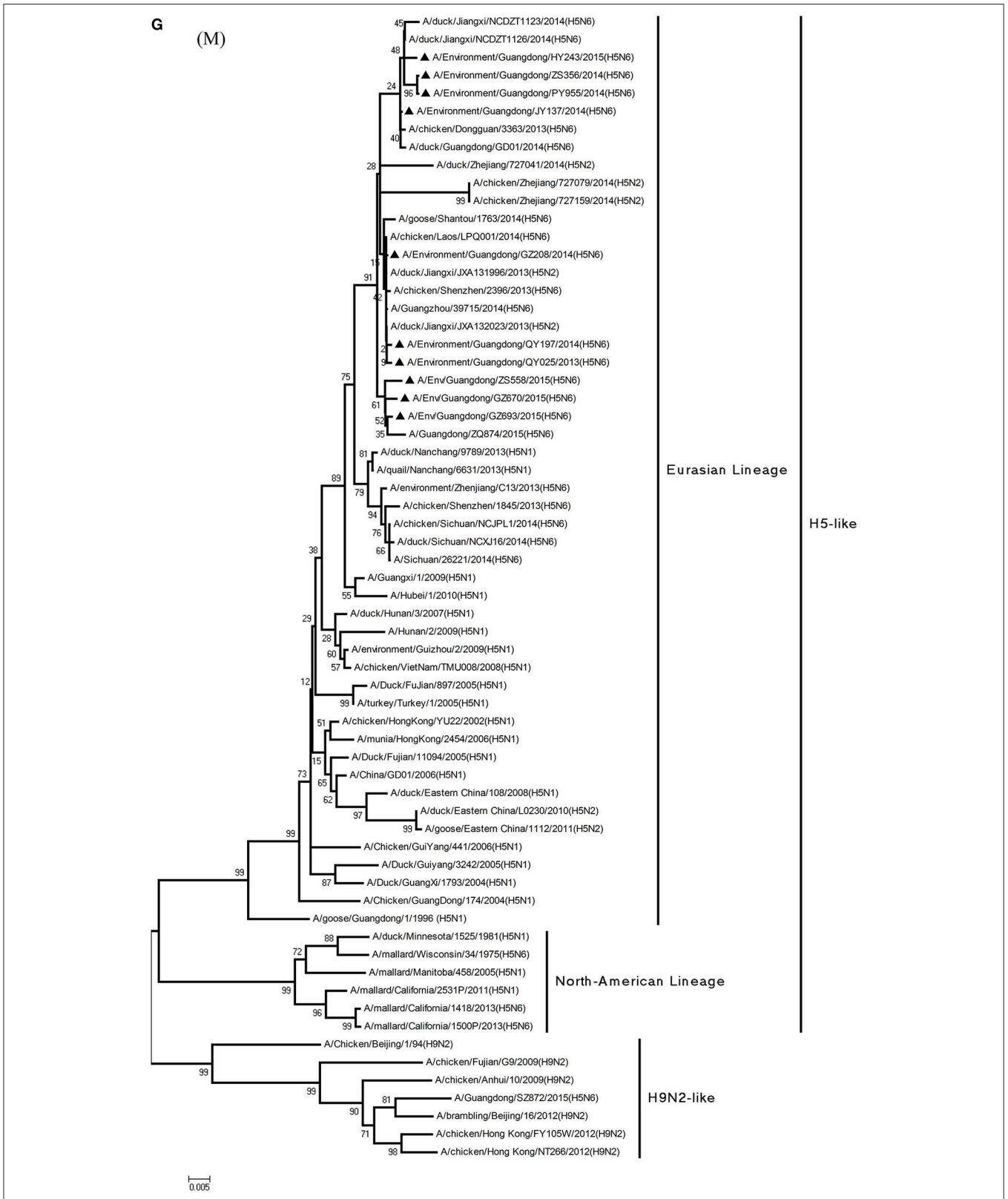


FIGURE 1 | Continued

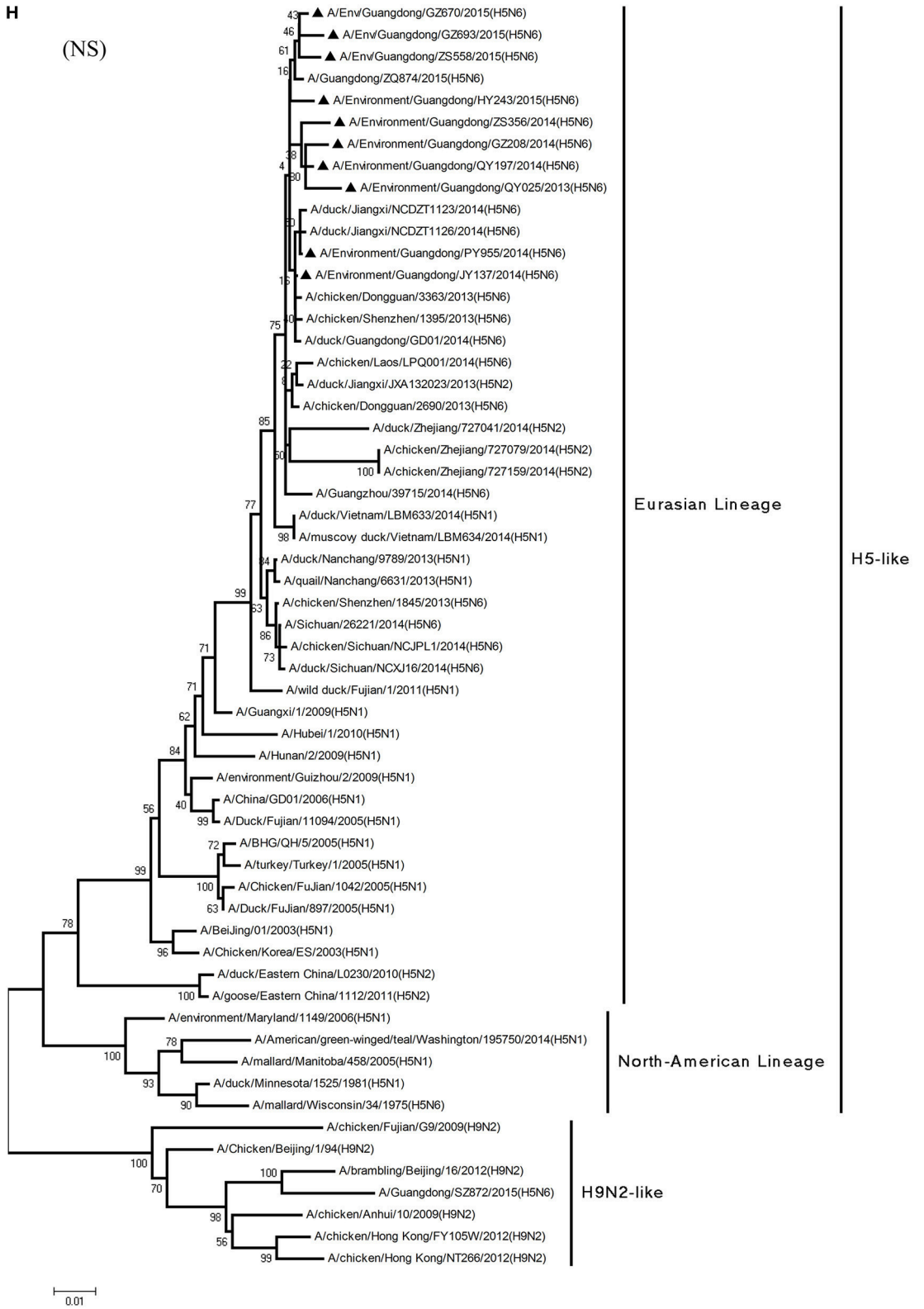


FIGURE 1 | Phylogenetic analyses of the open reading frames of H5N6-subtype avian influenza viruses. Viruses highlighted with black triangles (▲) were characterized in the present study. The tree was constructed using the neighbor-joining of Molecular Evolutionary Genetics Analysis 6.06, with 1000 bootstrap trials to ensure confidence in the groupings.

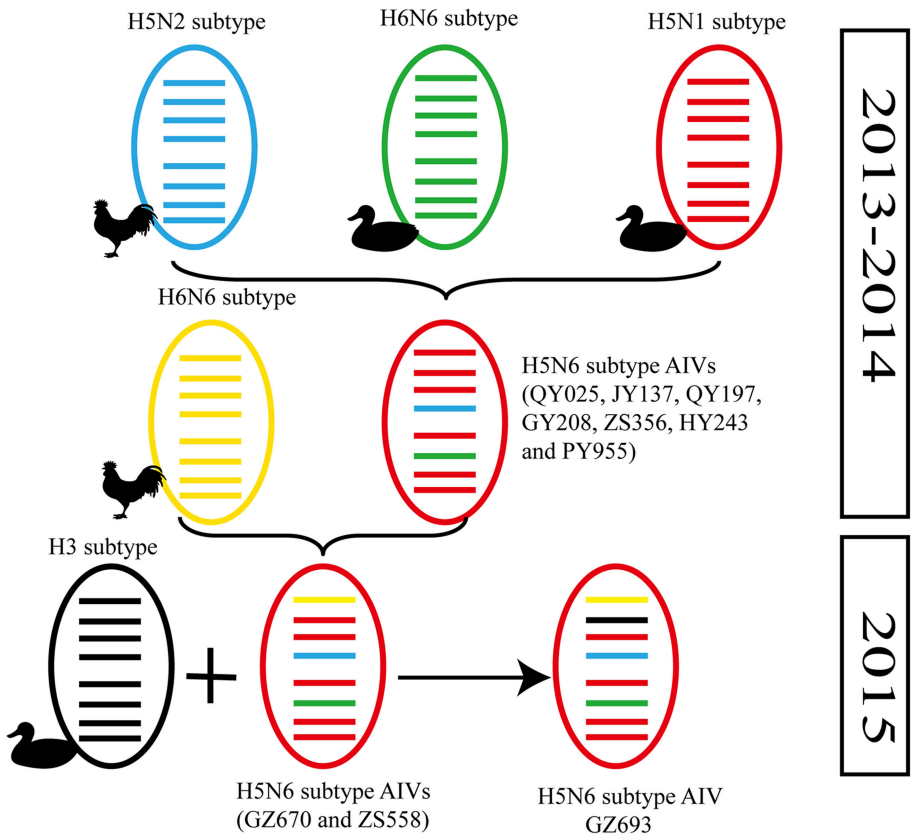


FIGURE 2 | Possible genetic contributions of known donor AIVs to the novel H5N6 avian influenza virus in Southern China. From top to bottom, the eight genes in each schematic virus particle are the PB2, PB1, PA, HA, NP, NA, M, and NS genes. Genes of the same lineage appear in the same color.

number) and position S31 of the M gene, showed no mutations (Scholtissek et al., 1998; Suzuki et al., 2003).

The PB2 gene of the 10 isolated viruses was E at position 627 and D at position 701, which indicates that all isolated viruses derived from avian sources (Li et al., 2005). At the same time, all environmental viruses were M at position 317 of the PB1 protein, which implies that they are hardly either pathogenic or non-pathogenic to mice (Katz et al., 2000). The AIVs could suppress a host's antiviral defenses relative to the antiviral effects of cytokines such as interferon. All viruses had P42S and D92E mutations in the NS1 protein, which suggests that they could enhance resistance to cytokines (Jiao et al., 2008; Qi et al., 2009).

DISCUSSION

At present, H5N1 AIVs have become endemic in waterfowl and domestic poultry in China, Southeast Asia, North America, and Africa, where they have evolved into multiple phylogenetic lineages (WHO/OIE/FAO, 2012). The regular transmission of H5N1 HPAIVs among waterfowl and domestic poultry has facilitated genetic diversity among circulating clades in poultry in China (Duan et al., 2008; Vijaykrishna et al., 2008). In particular, the AIVs of clades 2.3.2, 2.3.4, and 7.2 have cocirculated predominantly in domestic poultry and waterfowl in China continuously since 2007 (Smith et al., 2009; Jiang et al., 2010; Li et al., 2010). At the same time, evolutionary clades such as 2.3.4.5 and 2.3.4.6—recently redefined as clade 2.3.4.4—have been reported (Gu et al., 2013). Moreover, H5-subtype AIVs from clade 2.3.4.4 appear to be gradually replacing AIVs from clade 2.3.4.2, especially in waterfowl. In March 2014, an emergent H5N6 AIV caused an outbreak in poultry in Laos (Wong et al., 2015), and later, a flock of ducks was infected with H5N6 AIVs in Guangdong Province (Shen et al., 2015). Genetic analysis suggested that subtype H5N6 AIVs originated from clade 2.3.2.1b and variant clade 2.3.4 in H5N1 AIVs (Shen et al., 2015; Wong et al., 2015). As phylogenic analysis shows, of the 10 H5N6 AIVs isolated as part of LPM surveillance during 2013–2015, all environmental samples belonged to novel clade 2.3.4.4 and probably evolved to form a new subcluster, unlike those of H5N6 s previously identified in Sichuan Province.

Alongside HA evolution, the NA gene of H5N1 AIV has frequently reassorted with other subtypes of AIVs circulating in poultry (Zhao et al., 2008; Neumann et al., 2010). The new reassortments, including H5N3, H5N6, and H5N8, together with H7N9 and H9N2, are currently cocirculating in domestic poultry and waterfowl worldwide. In our study, H5N6 AIVs were natural recombinants, the NA gene of which derived from H6N6 AIVs circulating broadly in ducks in southern China. Within the first

7 months of 2014, internal genes of H5N6 reassortants were derived from the genetic backbone of the H5N1 subtype (Wu et al., 2015). Interestingly, for H5N6 viruses isolated after 2015, we noted the divergence of three H5N6 reassortants—namely, GZ670, GZ693, and ZS558—isolated after 2015 (**Figure 2**). The PB2 genes of GZ670, GZ693, and ZS558 were not grouped into the same clusters as other reported H5N6 viruses, but within the same clusters as H6N2 AIVs. Furthermore, the PB1 gene of GZ693 was clustered as a H3-subtype AIV. These results indicate that H5N6 AIV is constantly evolving, and as such, novel AIVs possessing H5- and H6-derived internal genes and other AIVs possessing specific mammal-derived mutations could enhance virulence and transmissibility in humans.

After December 2014, the first H5N6 AIV infections in humans in Guangdong Province seemed to an appeared to stop. From December 2015 to January 2016, however, five H5N6 AIV infections in humans were reported in Guangdong Province (WHO, 2014b, 2016a,b,c). Consistent with the evolution of H5N6 AIVs isolated from LPMs, the sequences of H5N6 AIVs isolated from patients are constantly evolving. The whole gene sequences of the first human H5N6 AIV were similar to those of the H5N6 AIVs isolated in early 2015 in LPMs in Guangdong Province. Meanwhile, the whole gene sequences of the other four human H5N6 AIVs were consistent with those of H5N6 AIVs isolated from LPMs in late 2015. Molecular characterization and phylogenetic analysis exhibited a highly close genetic relationship between the viruses isolated from humans and LPMs, thereby suggesting that infection in humans might be caused by the LPM environment.

LPMs have been deemed potential hotbeds for infection with H5N1 and H7N9 AIVs in humans (Wan et al., 2011; Shi et al., 2013). Some human–human transmission of AIVs (e.g., H5N1 and H7N9) has been reported (Wang et al., 2008; Qi et al., 2013), and as of February 2016, nine confirmed human infections with subtype H5N6 had occurred in China's Sichuan, Guangdong, and Yunnan Provinces (CDC China, 2014; WHO, 2015b,c, 2016a,b,c). In particular, the patient infected with H5N6 AIV in Guangzhou had visited an LPM before the onset of illness and could have acquired the infection there (Yang et al., 2015). The other patient infected with H5N6 AIV and who died in Sichuan Province was a merchant at a local LPM. Moreover, the other seven cases of infection had visited LPMs in the past. Perhaps above all, we isolated 10 H5N6 AIVs in LPMs, which indicates that LPMs are potential sources of AIV infection in humans.

In conclusion, we analyzed the evolution of H5N6 samples isolated from LPM environments. Epidemiological and experimental data suggest that the H5N6 subtype currently has a limited capacity for chicken–human or environment–human transmission. LPMs can provide sufficient opportunities for close contact among waterfowl, domestic poultry, mammals, and humans, as well as potential AIV infection, which in turn results in the emergence of novel AIVs. Large-scale surveillance of LPMs therefore continues to be essential to identifying novel reassortants and sequence mutations among existing AIV subtypes.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: RY, CK. Performed the experiments: RY, ZW, JW, LL. Analyzed the data: RY. Contributed reagents/materials/analysis tools: RY, LZ, YS, HN, JL, XZ, CK. Wrote the paper: RY, YK.

ACKNOWLEDGMENTS

This work was supported by grants from the Research Project of H7N9 Influenza of Guangdong (2014; No. 1046), the Scientific and Technological Research of Prevention and Control of H7N9 Subtype Avian Influenza Virus (20140224), the Science and Technology Planning Project of Guangzhou City, China (Grant Number. 2014J4100091, 2013J4200020), Science and Technology Planning Project of Guangdong Province (Grant Number. 2013B020307006), and the Medical Scientific Research Foundation of Guangdong Province, China (Grant Number. A2012078).

REFERENCES

CDC China (2014). *Avian Influenza A(H5N6) Virus was Detected in a Respiratory Tract Sample from a Patient Who Died of Severe Pneumonia in Sichuan.* Available online at: http://www.chinacdc.cn/mtdx/crbxx/201405/t20140507_96416.htm

Claas, E. C., Osterhaus, A. D., van Beek, R., De Jong, J. C., Rimmelzwaan, G. F., Senne, D. A., et al. (1998). Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* 351, 472–477. doi: 10.1016/S0140-6736(97)11212-0

Duan, L., Bahl, J., Smith, G. J., Wang, J., Vijaykrishna, D., Zhang, L. J., et al. (2008). The development and genetic diversity of H5N1 influenza virus in China, 1996-2006. *Virology* 380, 243–254. doi: 10.1016/j.virol.2008.07.038

Gu, M., Zhao, G., Zhao, K., Zhong, L., Huang, J., Wan, H., et al. (2013). Novel variants of clade 2.3.4 highly pathogenic avian influenza A(H5N1) viruses, China. *Emerg. Infect. Dis.* 19, 2021–2024. doi: 10.3201/eid1912.130340

Ha, Y., Stevens, D. J., Skehel, J. J., and Wiley, D. C. (2001). X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11181–11186. doi: 10.1073/pnas.201401198

Huang, K., Zhu, H., Fan, X., Wang, J., Cheung, C. L., Duan, L., et al. (2012). Establishment and lineage replacement of H6 influenza viruses in domestic ducks in southern China. *J. Virol.* 86, 6075–6083. doi: 10.1128/JVI.06389-11

Jiang, W. M., Liu, S., Chen, J., Hou, G. Y., Li, J. P., Cao, Y. F., et al. (2010). Molecular epidemiological surveys of H5 subtype highly pathogenic avian influenza viruses in poultry in China during 2007-2009. *J. Gen. Virol.* 91(Pt 10), 2491–2496. doi: 10.1099/vir.0.023168-0

Jiao, P., Tian, G., Li, Y., Deng, G., Jiang, Y., Liu, C., et al. (2008). A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. *J. Virol.* 82, 1146–1154. doi: 10.1128/JVI.01698-07

Katz, J. M., Lu, X., Tumpey, T. M., Smith, C. B., Shaw, M. W., and Subbarao, K. (2000). Molecular correlates of influenza A H5N1 virus pathogenesis in mice. *J. Virol.* 74, 10807–10810. doi: 10.1128/JVI.74.22.10807-10810.2000

Li, Y., Shi, J., Zhong, G., Deng, G., Tian, G., Ge, J., et al. (2010). Continued evolution of H5N1 influenza viruses in wild birds, domestic poultry,

and humans in China from 2004 to 2009. *J. Virol.* 84, 8389–8397. doi: 10.1128/JVI.00413-10

Li, Z., Chen, H., Jiao, P., Deng, G., Tian, G., Li, Y., et al. (2005). Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* 79, 12058–12064. doi: 10.1128/JVI.79.18.12058-12064.2005

Matsuoka, Y., Swayne, D. E., Thomas, C., Rameix-Welti, M. A., Naffakh, N., Warnes, C., et al. (2009). Neuraminidase stalk length and additional glycosylation of the hemagglutinin influence the virulence of influenza H5N1 viruses for mice. *J. Virol.* 83, 4704–4708. doi: 10.1128/JVI.01987-08

Neumann, G., Green, M. A., and Macken, C. A. (2010). Evolution of highly pathogenic avian H5N1 influenza viruses and the emergence of dominant variants. *J. Gen. Virol.* 91(Pt 8), 1984–1995. doi: 10.1099/vir.0.020750-0

OIE (2015). *Update on Highly Pathogenic Avian Influenza in Animals (Type H5 and H7)*. Available online at: http://www.oie.int/en/animal-health-in-the-world/update-on-avian-influenza/2015/

Qi, X., Li, X., Rider, P., Fan, W., Gu, H., Xu, L., et al. (2009). Molecular characterization of highly pathogenic H5N1 avian influenza A viruses isolated from raccoon dogs in China. *PLoS ONE* 4:e4682. doi: 10.1371/journal.pone.0004682

Qi, X., Qian, Y.-H., Bao, C.-J., Guo, X.-L., Cui, L.-B., Tang, F.-Y., et al. (2013). Probable person to person transmission of novel avian influenza A (H7N9) virus in Eastern China, 2013: epidemiological investigation. *BMJ* 347:f4752. doi: 10.1136/bmj.f4752

Scholtissek, C., Quack, G., Klenk, H. D., and Webster, R. G. (1998). How to overcome resistance of influenza A viruses against adamantane derivatives. *Antiviral Res.* 37, 83–95. doi: 10.1016/S0166-3542(97)00061-2

Shen, H., Wu, B., Chen, Y., Bi, Y., and Xie, Q. (2015). Influenza A(H5N6) virus reassortant, Southern China, 2014. *Emerg. Infect. Dis.* 21, 1261–1262. doi: 10.3201/eid2107.140838

Shi, J., Deng, G., Liu, P., Zhou, J., Guan, L., Li, W., et al. (2013). Isolation and characterization of H7N9 viruses from live poultry markets—implication of the source of current H7N9 infection in humans. *Chin. Sci. Bull.* 58, 1857–1863. doi: 10.1007/s11434-013-5873-4

Smith, G. J., Vijaykrishna, D., Ellis, T. M., Dyrting, K. C., Leung, Y. H., Bahl, J., et al. (2009). Characterization of avian influenza viruses A (H5N1) from wild birds, Hong Kong, 2004-2008. *Emerg. Infect. Dis.* 15, 402–407. doi: 10.3201/eid1503.081190

Subbarao, K., Klimov, A., Katz, J., Regnery, H., Lim, W., Hall, H., et al. (1998). Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* 279, 393–396. doi: 10.1126/science.279.5349.393

Suzuki, H., Saito, R., Masuda, H., Oshitani, H., Sato, M., and Sato, I. (2003). Emergence of amantadine-resistant influenza A viruses: epidemiological study. *J. Infect. Chemother.* 9, 195–200. doi: 10.1007/s10156-003-0262-6

Swayne, D. E., and Halvorson, D. A., (2003). "Influenza," in *Diseases of Poultry*, eds Y. M. Saif, H. J. Barnes, J. R. Glisson, A. M. Fadly, L. R. McDougald, and D. E. Swayne (Ames, IA: Iowa State University Press), 135–160.

Vijaykrishna, D., Bahl, J., Riley, S., Duan, L., Zhang, J. X., Chen, H., et al. (2008). Evolutionary dynamics and emergence of panzootic H5N1 influenza viruses. *PLoS Pathog.* 4:e1000161. doi: 10.1371/journal.ppat.1000161

Wan, X. F., Dong, L., Lan, Y., Long, L. P., Xu, C., Zou, S., et al. (2011). Indications that live poultry markets are a major source of human H5N1 influenza virus infection in China. *J. Virol.* 85, 13432–13438. doi: 10.1128/JVI.05266-11

Wang, H., Feng, Z., Shu, Y., Yu, H., Zhou, L., Zu, R., et al. (2008). Probable limited person-to-person transmission of highly pathogenic avian influenza A (H5N1) virus in China. *Lancet* 371, 1427–1434. doi: 10.1016/S0140-6736(08)60493-6

WHO/OIE/FAO (2012). Continued evolution of highly pathogenic avian influenza A (H5N1): updated nomenclature. *Influenza Other Respir. Viruses* 6, 1–5. doi: 10.1111/j.1750-2659.2011.00298.x

WHO (2014a). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/28-december-2014-avian-influenza/en/

WHO (2014b). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/28-december-2014-avian-influenza/en/

WHO (2014c). *Influenza at the Human-Animal Interface (HAI)*. Available online at: http://www.who.int/influenza/human_animal_interface/en/

WHO (2014d). *WHO China Statement on H5N6*. Available online at: http://www.wpro.who.int/china/mediacentre/releases/2014/20140507/en/

WHO (2015a). *Cumulative Number of Confirmed Human Cases for Avian Influenza A(H5N1) Reported to WHO, 2003-2015*. Available online at: http://www.who.int/influenza/human_animal_interface/EN_GIP_20150717cumulativeNumberH5N1cases.pdf?ua=1

WHO (2015b). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/12-february-2015-avian-influenza/en/

WHO (2015c). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/14-july-2015-avian-influenza/en/

WHO (2016a). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/4-january-2016-avian-influenza-china/en/

WHO (2016b). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/11-january-2016-avian-influenza-china/en/

WHO (2016c). *Human Infection with Avian Influenza A(H5N6) Virus – China*. Available online at: http://www.who.int/csr/don/26-january-2016-avian-influenza-china/en/

Wong, F. Y., Phommachanh, P., Kalpravidh, W., Chanthavisouk, C., Gilbert, J., Bingham, J., et al. (2015). Reassortant highly pathogenic influenza A(H5N6) virus in Laos. *Emerg. Infect. Dis.* 21, 511–516. doi: 10.3201/eid2103.141488

Wu, H., Lu, R., Peng, X., Xu, L., Cheng, L., Lu, X., et al. (2015). Novel reassortant highly pathogenic H5N6 avian influenza viruses in poultry in China. *Infect. Genet. Evol.* 31, 64–67. doi: 10.1016/j.meegid.2015.01.019

Yang, Z. F., Mok, C. K., Peiris, J. S., and Zhong, N. S. (2015). Human infection with a novel avian Influenza A(H5N6) virus. *N. Engl. J. Med.* 373, 487–489. doi: 10.1056/NEJMc1502983

Yuan, R., Cui, J., Zhang, S., Cao, L., Liu, X., Kang, Y., et al. (2014). Pathogenicity and transmission of H5N1 avian influenza viruses in different birds. *Vet. Microbiol.* 168, 50–59. doi: 10.1016/j.vetmic.2013.10.013

Yuan, R., Zou, L., Kang, Y., Wu, J., Zeng, X., Lu, J., et al. (2016). Reassortment of avian influenza A/H6N6 viruses from live poultry markets in Guangdong, China. *Front. Microbiol.* 7:65. doi: 10.3389/fmicb.2016.00065

Zhao, Z. M., Shortridge, K. F., Garcia, M., Guan, Y., and Wan, X. F. (2008). Genotypic diversity of H5N1 highly pathogenic avian influenza viruses. *J. Gen. Virol.* 89(Pt 9), 2182–2193. doi: 10.1099/vir.0.2008/001875-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Yuan, Wang, Kang, Wu, Zou, Liang, Song, Zhang, Ni, Lin and Ke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 16 November 2016
doi: 10.3389/fmicb.2016.01766

Novel H7N2 and H5N6 Avian Influenza A Viruses in Sentinel Chickens: A Sentinel Chicken Surveillance Study

Teng Zhao[1†], Yan-Hua Qian[2†], Shan-Hui Chen[2], Guo-Lin Wang[1], Meng-Na Wu[1], Yong Huang[1], Guang-Yuan Ma[2], Li-Qun Fang[1], Gregory C. Gray[3], Bing Lu[2], Yi-Gang Tong[1], Mai-Juan Ma[1] and Wu-Chun Cao[1*]*

[1] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China, [2] Wuxi Center for Disease Control and Prevention, Wuxi, China, [3] Division of Infectious Diseases, Global Health Institute, Nicholas School of the Environment, Duke University, Duke University Medical Center, Durham, NC, USA

OPEN ACCESS

Edited by:
José A. Melero,
Instituto de Salud Carlos III, Spain

Reviewed by:
Jianwei Wang,
China Academy of Chinese Medical Sciences, China
Shailesh D. Pawar,
National Institute of Virology, India

***Correspondence:**
Mai-Juan Ma
mjma@163.com
Wu-Chun Cao
caowc@bmi.ac.cn

[†]These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *05 July 2016*
Accepted: *20 October 2016*
Published: *16 November 2016*

Citation:
Zhao T, Qian Y-H, Chen S-H, Wang G-L, Wu M-N, Huang Y, Ma G-Y, Fang L-Q, Gray GC, Lu B, Tong Y-G, Ma M-J and Cao W-C (2016) Novel H7N2 and H5N6 Avian Influenza A Viruses in Sentinel Chickens: A Sentinel Chicken Surveillance Study. Front. Microbiol. 7:1766. doi: 10.3389/fmicb.2016.01766

In 2014, a sentinel chicken surveillance for avian influenza viruses was conducted in aquatic bird habitat near Wuxi City, Jiangsu Province, China. Two H7N2, one H5N6, and two H9N2 viruses were isolated. Sequence analysis revealed that the H7N2 virus is a novel reassortant of H7N9 and H9N2 viruses and H5N6 virus is a reassortant of H5N1 clade 2.3.4 and H6N6 viruses. Substitutions V186 and L226 (H3 numbering) in the hemagglutinin (HA) gene protein was found in two H7N2 viruses but not in the H5N6 virus. Two A138 and A160 mutations were identified in the HA gene protein of all three viruses but a P128 mutation was only observed in the H5N6 virus. A deletion of 3 and 11 amino acids in the neuraminidase stalk region was found in two H7N2 and H5N6 viruses, respectively. Moreover, a mutation of N31 in M2 protein was observed in both two H7N2 viruses. High similarity of these isolated viruses to viruses previously identified among poultry and humans, suggests that peridomestic aquatic birds may play a role in sustaining novel virus transmission. Therefore, continued surveillance is needed to monitor these avian influenza viruses in wild bird and domestic poultry that may pose a threat to poultry and human health.

Keywords: avian influenza A virus, H7N2 virus, H9N2 virus, H5N6 virus, sentinel chicken, transmission

INTRODUCTION

Since the first emergence of avian influenza A(H7N9) virus in early 2013 (Gao et al., 2013), as of January 24, 2016, China has experienced four epidemic waves resulting in 711 laboratory-confirmed human infections with 283 deaths (Chinese National Influenza Center, 2016). The continuing circulation and evolution of H7N9 viruses in poultry (Lam et al., 2015), and the increasing number of human infections clearly identify H7N9 virus as an ongoing public health threat.

Domestic poultry are considered the main reservoir for H7N9 virus that cause human infections (Yu et al., 2013). However, wild birds, such as tree sparrows, songbirds, parakeets, and finches, have also been implicated as a source of virus transmission (Jones et al., 2014, 2015; Zhao et al., 2014). Because these wild birds share common space and resources with wild migratory birds, poultry, and humans, it is conceivable that wild avian species, especially aquatic birds, songbirds, passerine, and other small terrestrial birds might be infected with H7N9 virus and serve

as vectors for dissemination of the virus to domestic poultry. However, research regarding H7N9 virus transmission from wild birds to poultry or vice versa has been sparse. Experimental studies of H7N9 virus natural transmission from wild avian birds to poultry may be necessary to examine such hypotheses. In view of this, surveillance of sentinel chicken for avian influenza viruses in a peridomestic aquatic areas (Coman et al., 2014) was undertaken to study transmission of H7N9 and other subtypes of avian influenza viruses.

MATERIALS AND METHODS

Study Site

During the period from January to October of 2014, a sentinel chicken surveillance study was conducted in Taihu Lake (**Figure 1A**) in Wuxi City, Jiangsu Province, China, where 13 human infections with H7N9 virus have been reported. Taihu Lake (30°55′40″–31°32′58″ N, 119°52′32″–120°36′10″ E) is China's third largest freshwater lake with an area of about 2338 km^2. It is located in the Yangtze River delta and crosses Wuxi City, where 40 species of wild birds and 19 species of resident birds can be found and most of birds are winter migration birds (Zhao et al., 1990). These mainly include little egrets, herons, green-backed herons, black-crowned night herons, sparrows, quails, Chinese hwamei, yellow-browed warblers, Eurasian magpies, azure-winged magpies, and house swallows. The sentinel site was located on a foothill at the foot of the Lake's Mashan Mountain (**Figure 1B**) and was about 100 m^2 in area such that the sentinel chickens would be free to move.

Sentinel Chickens and Sample Collection

We employed chickens as the sentinel bird for surveillance of avian influenza virus because chicken was one of main reservoir of H7N9 virus. In addition, influenza A virus negative chickens were easier to obtain than other domestic poultry such as ducks. A chicken sample size of 126 was chosen in order to detect a 10% avian influenza virus prevalence with a 95% confidence interval of ±5% with 80% power. However, a total of 135 chickens were used considering the death of chickens during the surveillance. On January 6, 2014, coincided with wild bird migration into the Taihu Lake, a total of 135 6-month-old, specific-pathogen-free (SPF) white leghorn chickens (Beijing Merial Vital Laboratory Animal Technology Co., Ltd., Beijing, China) were placed at sentinel surveillance site. A high rate of females to males (4:1) was chosen to avoid fighting among roosters. Before the sentinel chickens were placed into the surveillance site, cloacal and throat swabs of each chicken were collected, pooled, and screened for influenza A virus using a real-time RT-PCR assay which targeted the matrix gene (World Health Organization, 2014). In addition, an electronica ring (Beijing Raybaca Technology Co., Ltd) was fixed at the one leg of each chicken to identify their unique identity. The chickens were cared for and fed by a local farmer who agreed to assist us in this surveillance study. Each evening the chickens were encouraged with food to return to their protective enclosure.

During the surveillance period, cloacal and throat swabs from each chicken were collected and pooled together as one swab sample each month from February 6 to April 3 and from July 31 to August 20, and weekly from April 9 to June 26 as the number of wild birds increased in number due to in-migrations. All samples were placed in transport medium consisting of phosphate-buffered saline (PBS) containing 50% glycerol, penicillin (2000 U/ml), gentamicin (250 mg/ml), polymixin B (2000 U/ml), nystatin (500 U/ml), ofloxacin HCl (60 mg/ml), and sulfamethoxazole (200 mg/ ml) and were kept at 4°C for ∼4 h

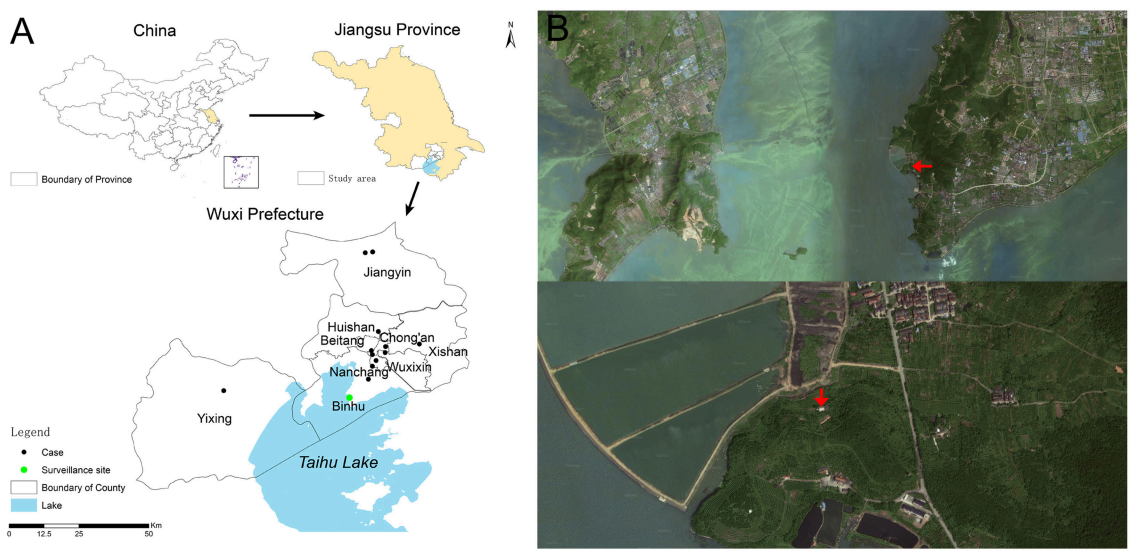


FIGURE 1 | Sentinel surveillance study site, Wuxi City, Jiangsu Province of China. (A) Location of surveillance study site and distribution of human infections with influenza A (H7N9) virus; **(B)** Satellite map of study site, red arrows indicate study site.

until they were transported to the laboratory, where they were stored at −80°C for virus isolation within 2 days.

The chicken handling and sampling was performed in accordance with experimental animal administration and the ethics committee of Academy of Military of Medical Sciences. The protocol was approved by the biosafety committee of Beijing institute of Microbiology and Epidemiology.

Virus Isolation and Sequencing Analysis

All of swab samples were inoculated into the allantoic cavities of 9-day-old SPF embryonated chicken eggs (Beijing Merial Vital Laboratory Animal Technology Co., Ltd., Beijing, China) for virus isolation. Allantoic fluid were harvested after incubation at 37°C for 72 h or immediately after embryos died, and tested for hemagglutination (HA) activity with horse red blood cell and real time RT-PCR to detect the presence of influenza A virus. All virus isolation procedures were conducted in a biosafety level 3 facility. The whole viral genome of isolated viruses were sequenced by using Ion Torrent PGM sequencing technology (Life Technologies, Grand Island, NY, USA).

Plaque Purification of Two Co-infected Samples

Two swab samples, #4315 and #6395, that contained mixed H7 and H9 subtypes were subjected to the standard plaque purification (Tobita et al., 1975) with a slight modification to segregate clones of individual viruses. Briefly, 100 μl aliquots of serial 10-fold dilutions of swab sample were inoculated into confluent monolayer MDCK cells cultured in 6 well-plates, MDCK cells were incubated at 37°C in a CO_2 incubator for 1 h for virus absorption. Then covered with 1% agar containing MEM, penicillin (1000 unit/ml), streptomycin (1000 μg/ml), fungizone (2.5 μg/ml), MEM vitamin solution, 0.5% bovine serum albumin, and 2 μg/ml trypsin. After 2 days of incubation at 37°C in a CO_2 incubator and stained with 0.025% neutral red containing 1% agarose, 42 and 53 plaques for #4315 and #6395 were initially picked for further purification and resuspended in 500 μl of infection medium. The subtypes of all resuspended virus stocks were verified using real time RT-PCR analysis (rRT-PCR), and virus stocks showing a single subtype were further purified in SPF chicken embryos. After three serial rounds of purification, the H7N2 and H9N2 virus isolates were successfully separated and their genome were fully sequenced. The sequences data generated in this study were deposited in Global Initiative on Sharing All Influenza Data (accession nos. EPI_ISL_223199–EPI_ISL_223202, EPI_ISL_223364). We used the maximum likelihood method to generate phylogenetic trees using MEGA version 6.0.6 (www.megasoftware.net/) with a bootstrapping resampling process (1000 replications) to assess the robustness of individual nodes of the phylogeny.

RESULTS

A total of 1623 pooled samples were collected from the sentinel chickens. In total, three (0.18%) of 1623 samples (samples #4315, #6395, and #7393) were positive for HA activity after virus isolation. Subtype determination, using rRT-PCR analysis targeting the H5, H7, and H9 HA genes, showed that specimens #4315 and #6395 were positive for all H7 and H9 genes, and specimen #7393 was only positive for H5 gene. To confirm these results, three HA-positive allantoic fluid specimens were subjected to sequencing on Ion Torrent PGM, and results consist with rRT-PCR that both specimens #4315 and #6395 that contained influenza A H7, H9, and N2 gene segments suggesting a coinfection of H7 and H9 subtypes, while specimens #7393 only presented with H5 and N6 type of influenza A virus. We then attempted to isolate individual H7 and H9 viruses from these two samples using a plaque purification. Fortunately, we were able to segregate the expected subtypes, H7N2 and H9N2 from two co-infected samples. Thus, totally five strains were isolated and designated as A/chicken/Wuxi/SC4315/2014 (H7N2), A/chicken/Wuxi/SC6395/2014 (H7N2), A/chicken/Wuxi/SC4315/2014 (H9N2), A/chicken/Wuxi/SC6395/2014 (H9N2), and A/chicken/Wuxi/SC7393/2014 (H5N6).

To characterize H7N2 viruses, molecular and phylogenetic analysis were conducted. The HA gene segments of the two H7N2 viruses most closely matched those of H7N9 viruses isolated from human and poultry in China in 2013 with an up to 99.6% nucleotide sequence identity (**Table 1**). Their NA segments were closely related to that of an H9N2 virus isolated from a chicken in Jiangsu Province, China during 2013 (99.5% identity). Their other six internal genes were derived from H9N2 viruses currently found in China with identities range from 99.5–99.7% (**Table 1**). In comparison with two H9N2 viruses in present study, all gene segments of two H7N2 viruses showed 99.0–100% total nucleotide sequence identity (**Table 1**). The two H7N9 viruses showed 100% total nucleotide sequence identity for HA, NP, NS, and PA, and >99.5% identity for other gene segments. Phylogenetic analysis based on HA genes showed that these two H7N2 viruses closely related with those H7N9 isolated in first and second waves, but were form an independent clade (bootstrap support of 99%) from a second H7N9 virus wave in China's Zhejiang, Jiangsu, and Jiangxi Provinces (**Figure 2**). Whereas, the other seven gene segments, including NA, clustered with those of the H9N2 viruses circulated in China and two H9N2 viruses identified in present study (**Figure 2** and **Supplementary Figure 1**). These results suggest that the two H7N2 viruses are novel reassortant of H7N9 and H9N2 viruses recently circulated in China.

As for that H5N6 virus, the HA gene phylogeny showed that the virus was closely related to A/chicken/Dongguan/4259/2013(H5N6) clade 2.3.4.4 from Guangdong Province, China. The virus likely originated from highly pathogenic H5 avian influenza viruses with N1, N2, and N8 neuriminidases detected in poultry in China since 2010 and from viruses detected in Vietnam in 2014 (**Figure 3**). The NA gene segment of the H5N6 virus clustered into Jiangxi lineage and likely originated from H6N6 avian influenza viruses circulating in China's Fujian and Guangdong Provinces (**Figure 3**). The six internal genes of the H5N6 virus were found to cluster with H5N6 strains identified in China and Vietnam (**Supplementary Figure 2**).

Molecular analysis of critical and apparent amino acid residues that may be associated with adaptation of an avian

TABLE 1 | Influenza A viruses with greatest nucleotide sequence identity to avian influenza A (H7N2) viruses and sequence comparison of isolated in present study from Wuxi, China, 2014.

Gene	Virus designation	Identity, %			
		1	2	3	4
A/CHICKEN/WUXI/SC63951/2014 (H7N2)					
HA	A/Anhui/1-BALF_RG8/2013(H7N9)	99.6	–	–	100.0
NA	A/chicken/Jiangsu/WX0204/2013(H9N2)	99.5	99.9	99.9	99.9
PB2	A/chicken/Beijing/16/2013(H9N2)	99.3	99.6	99.6	99.8
	A/chicken/Rizhao/55/2013(H9N2)				
PB1	A/chicken/Shandong/qd0917/2013(H9N2)	99.4	99.8	99.8	99.5
PA	A/chicken/Shandong/qd0917/2013(H9N2)	99.7	100.0	100.0	100.0
NP	A/chicken/Qingdao/2144/2013(H9N2)	99.5	99.9	99.9	100.0
M	A/chicken/Beijing/0331/2013(H9N2)	99.7	99.9	99.9	99.9
NS	A/chicken/Shandong/qd0917/2013(H9N2)	98.4	99.0	99.0	100.0
A/CHICKEN/WUXI/SC43151/2014 (H7N2)					
HA	A/Anhui/1-BALF_RG8/2013(H7N9)	99.6	–	–	–
NA	A/chicken/Jiangsu/WX0204/2013(H9N2)	99.6	100.0	100.0	–
PB2	A/chicken/Beijing/16/2013(H9N2)	99.3	99.6	99.6	–
	A/chicken/Rizhao/55/2013(H9N2)				
PB1	A/chicken/Shandong/qd0917/2013(H9N2)	98.9	99.2	99.2	–
PA	A/chicken/Shandong/qd0917/2013(H9N2)	99.7	100.0	100.0	–
NP	A/chicken/Qingdao/2144/2013(H9N2)	99.5	99.9	99.9	–
M	A/chicken/Beijing/0331/2013(H9N2)	99.8	100.0	100.0	–
NS	A/chicken/Shandong/qd0917/2013(H9N2)	98.4	99.0	99.0	–

1, identity determined by a BLAST search of the Influenza Sequence Database; 2, identity compared with A/chicken/Wuxi/SC6395/2014(H9N2); 3, identity compared with A/chicken/Wuxi/SC4315/2014(H9N2); 4, identity compared each other.

HA, hemagglutinin; NA, neuraminidase; PB, polymerase basic; PA, polymerase basic; NP, nucleoprotein; M, matrix; NS, nonstructural.

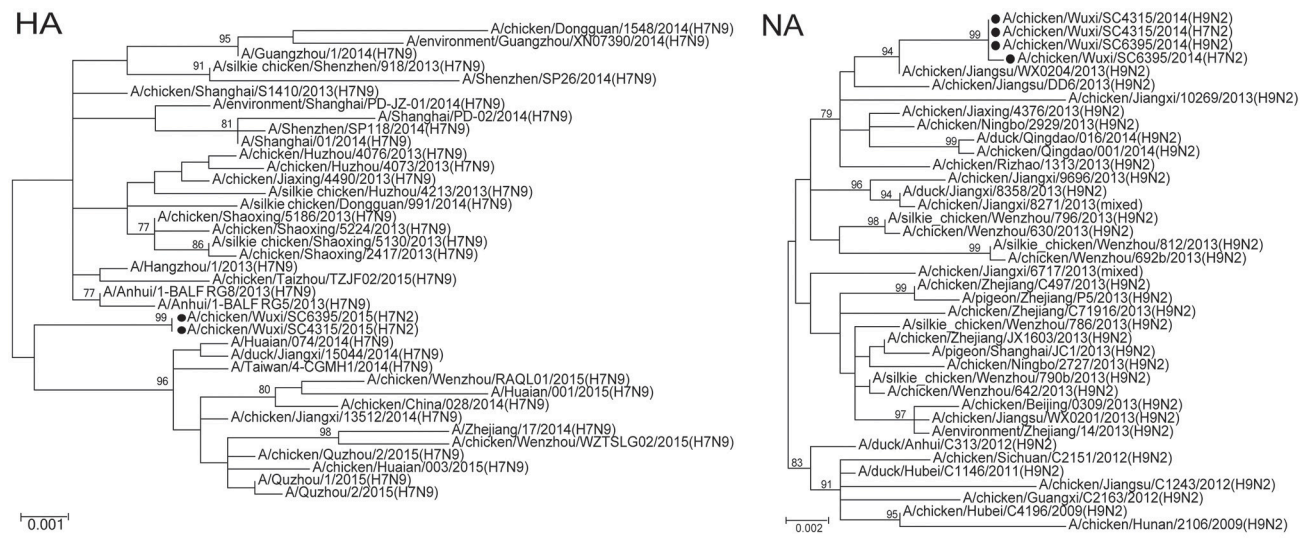


FIGURE 2 | Phylogenetic tree of hemagglutinin (HA) and neuraminidase (NA) genes of influenza A (H7N2) viruses isolated from sentinel chickens, Wuxi City, Jiangsu Province of China, 2014. Supporting bootstrap values >75 are shown. The black circle dot indicates the viruses isolated in present study.

virus to the mammalian host, virulence, and antiviral resistance of H7N2 and H5N6 viruses were conducted (**Table 2**). The HA gene protein of both two H7N2 viruses had a single basic amino acid (PEIPKGR↓G) at the cleavage site, indicating low pathogenic effects in poultry. Substitution at A138, A160, V186, and L226 (H3 numbering) was observed suggesting a possible increased affinity for binding human α2, 6-linked sialic acid receptors. Three amino acid deletions (63–65) were observed in

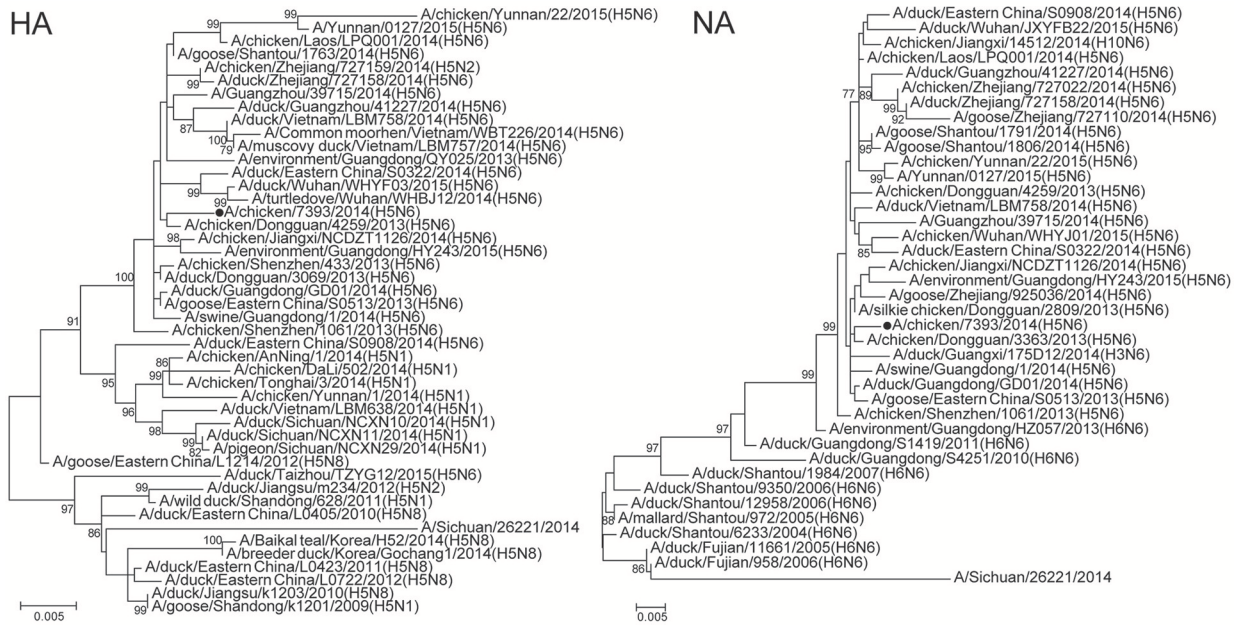


FIGURE 3 | Phylogenetic tree of hemagglutinin (HA) and neuraminidase (NA) genes of influenza A (H5N6) viruses isolated from sentinel chickens, Wuxi City, Jiangsu Province of China, 2014. Supporting bootstrap values >75 are shown. The black circle dot indicates the viruses isolated in present study.

the stalk region of NA protein. Although no substitutions of the H274K and R292K (N2 numbering) amino acid were observed, indicating that the viruses should still be sensitive to oseltamivir, zanamivir, and peramivir, mutation of V222 and K249 were observed in two viruses. Whereas, the HA gene protein of the H5N6 virus processed multiple basic amino acid motifs (RERRRKR↓G) at cleavage site, indicated highly pathogenic phenotype (**Table 2**). While the HA gene protein of H5N6 virus had the amino acids Q226 and G228 (H3 numbering) indicated that the virus preferentially bind to avian-like receptors (Stevens et al., 2006), but the mutation of P128, A138, and A160 may enhance binding capacity to human like receptors. The NA stalk of H5N6 virus possessed a deletion of 11 aa residues at positions 58–68 (N6 numbering), which suggested that the virus isolate might have different adaptation and virulence characteristics in poultry and mammals (Bi et al., 2015). Moreover, H5N6 virus contained a truncated PB1-F2 protein of 57 aa in length, which might influence their virulence in mammals (Zamarin et al., 2006). Amino acids V89, E627, and D701 were observed in the PB2 protein of H7N2 and H5N6 viruses, suggesting that these viruses were not yet fully adapted to infect mammals (Li et al., 2005). No drug resistance-associated mutations were observed in the H5N6 isolate but mutation N31 in M2 protein was found in both two H7N2 viruses. Other possible molecular markers for two subtype viruses that may associated with virulence are shown in **Table 2**.

DISCUSSION

We identified the novel H7N2 and H5N6 viruses during the surveillance of avian influenza virus using sentinel chickens in China in 2014. Although, we did not collect the samples from wild birds in local area because of restriction of regulations, our results suggest that the viruses were likely transmitted from wild birds to sentinel chickens in this aquatic environment because influenza A virus negative SPF chickens were used. However, we are uncertain which birds may have transmitted virus to the sentinel chickens or how that transmission occurred. While aquatic birds would be the prime transmission suspects, resident or passerine birds might also be transmitters (Jones et al., 2014). Avian influenza virus transmission could have occurred via another source such as contact with man, via bird feed, or contaminated drinking water (Jones et al., 2015). It is interesting that wild birds were observed mixing with sentinel chickens together, and they drank from the same water and ate the same bird feed (**Figure 4**). Therefore, most likely route for sentinel chickens infected with H7N2 and H5N6 viruses from wild bird is by means of drinking the water and eating the feed. However, the evidence for virus transmission via drinking water or eating the feed was limited because we did not collect these samples to test the presence of avian influenza virus.

Although, we did not find any H7N9 viruses among the sentinel chickens during the surveillance period, we identified the novel two H7N2 and one H5N6 viruses. Phylogenetic analysis revealed that the HA gene of H7N2 viruses were closely related to those of the H7N9 viruses, that emerged in China recently and the other seven gene segments (PB2, PB1, PA, NP, M, NS, and NA) were closely related to those of H9N2 viruses that were isolated from poultry in China, providing the evidence that H7N9 viruses continue to evolve and reassort with H9N2 viruses in poultry in China. Molecular analysis showed that the substitutions at A138, A160, V186, and L226 (H3 numbering) in

TABLE 2 | Influenza A viruses isolated in present study and mutation analysis of critical amino acid residues associated with virulence, adaptation to mammals or antiviral resistance, Wuxi, China, 2014.

	A/chicken/Wuxi/SC6395/2014	A/chicken/Wuxi/SC4315/2014	A/chicken/Wuxi/SC7393/2014	Comments
Status of chickens	Healthy	Healthy	Sick	
Collection date	May 7, 2014	June 19, 2014	August 20, 2014	
HA				
Cleavage	PEIPKGR↓G	PEIPKGR↓G	RERRRKR↓G	
S128P	S	S	**P**	Mammalian adaptation
A135T	A	A	V	
S138A	**A**	**A**	**A**	
N158D	N	N	N	
T160A	**A**	**A**	**A**	
G186V	**V**	**V**	N	
Q226L	**L**	**L**	Q	
G228S	G	G	G	
NA				
Deletion	63-65 (N2 numbering)	63-65 (N2 numbering)	58-68 (N6 numbering)	Enhance virus adaptation to domestic fowl
E119V	E	E	E	Oseltamivir resistance
D198N	D	D	**N**	
I222V	**V**	**V**	I	
H274Y	H	H	H	
R292K	R	R	R	
N294S	N	N	N	
Q136K	Q	Q	Q	Zanamivir resistance
D151E	D	D	D	
R224K	R	R	R	
R249K	**K**	**K**	**K**	
E276D	E	E	E	
S315N	S	S	K	
R371K	R	R	R	
PB2				
E158G	E	E	E	Increased the morbidity and mortality of H5 viruses in mice
T271A	T	T	T	Enhance polymerase activity in human cells
L89V	V	V	V	Mammalian host adaption and increased virulence in mice
E627K	E	E	E	
D701N	D	D	D	
PB1				
H99Y	H	H	H	H5 virus transmissible among ferrets
I368V	**V**	**V**	I	
PB1-F2				
S12Stop	S	S	S	Enhanced transmission and pathogenicity to human for pH1N1
N66S	N	N	n/a	Increased the virulence of an H5N1 virus
58–90 truncated	No	No	Yes	Increased virulence in mammals
PA				
T97I	T	T	T	Increased pathogenicity in mice of H5N2 virus
T85I	T	T	T	Involved host adaptation of pH1N1

(Continued)

TABLE 2 | Continued

	A/chicken/Wuxi/SC6395/2014	A/chicken/Wuxi/SC4315/2014	A/chicken/Wuxi/SC7393/2014	Comments
G186S	G	G	G	
L336M	L	L	L	
P400L	P	P	P	Enhanced polymerase activity in human cells
M423I	V	V	**I**	
V476A	**A**	**A**	**A**	
T552S	T	T	T	
NP				
A184K	**K**	**K**	**K**	Increased replication in mammalian cells pathogenicity
N319K	N	N	N	
M1				
N30D	D	D	D	Increased pathogenicity in mice of H5N1
T215A	A	A	A	
M2				
S31N	**N**	**N**	S	Antiviral resistance (amantadine)
NS1				
P42S	S	S	S	Increased virulence in mice or in pigs
D92E	D	D	**E**	

HA, hemagglutinin; NA, neuraminidase; PB, polymerase basic; PA, polymerase basic; NP, nucleoprotein; M, matrix; NS, nonstructural. Bold represents substation.



FIGURE 4 | Monitors captured partial image of birds during the surveillance period. Red arrows indicate observed birds in study site.

HA protein were observed, which is proposed to enhance binding capacity of human like receptors and mammalian adaption. A deletion of three amino acids in the viral NA stalk has been observed. It has been suggested that amino acid deletion within the stalk domain may be associated with transmission of influenza from wild birds to domestic poultry (Cauldwell et al., 2014). However, it is unclear whether such H7N2 viruses would be also detected in wild birds because samples from wild bird were not collected in local. The viruses should still be sensitive to oseltamivir, zanamivir, and peramivir as no substitutions

of the K274 and K292 amino acid were observed. Although, the single mutation of V222 or K249 may not be associated with marginal levels of oseltamivir resistance, increased drug resistance levels in association with other mutation (Simon et al., 2011). In addition, mutation of N31 in M1 protein suggests that the virus had the resistance to amantadine. Previous studies have showed that H7N9 isolates from humans with the mutation of glutamic acid to lysine at position 627 (E627K) in PB2 were much more efficient and more lethal in mice than H7N9 isolates from birds (Zhang et al., 2013, 2014). The H7N2 viruses identified in present study does not have the PB2 K627 and N701 mutations but had an V89 mutation suggests that the viruses have not yet fully adapted to infect mammals (Li et al., 2005). However, H7N2 virus identified by Shi et al. (2014) without K627 mutation but showed comparable replication ability in the lungs of mice with human H7N9 virus and a significantly higher replication than that of avian H7N9 virus. These findings suggest that the continued circulation of H7 viruses in nature will enable them to acquire more mutations or new gene constellations that might increase their virulence in animals or humans. Taken together these findings, the continued circulation of H7N9 and H9N2 viruses in poultry and persistent human infection of H7N9 virus, H7 viruses still pose seriously threats to human health, especially if they acquire mutations or new gene reassortments that increase their virulence in animals or humans.

Since the first H5N6 virus infections in human reported in China, a total of 15 patients have been reported. Previous studies have shown that H5N6 viruses isolated from the patients had similar identity with viruses isolated from the live poultry market and they are constantly evolving. In our study, we isolated a highly pathogenic avian influenza A (H5N6) virus that belong to clade 2.3.4.4 from a sentinel chicken. However, any death bird was not observed in study site during period of surveillance although H5N6 viruses have been detected in fatal wild birds (Yu et al., 2015). Sequence analysis revealed that the virus had a very closely genetic relationship with H5N6 viruses isolated from the patients and live poultry, thereby suggesting that direct contact with poultry contributed to the human infection. Furthermore, the H5N6 virus has distinct evolutionary characteristics, such as the P128, A138, and A160 mutations in the HA protein and an 11-aa deletion in the NA stalk, which may enhance their adaptation and infectivity in mammals, including humans. The virus may resistant to oseltamivir and zanamivir (N198 and K249) but sensitive to amantadine (S31). Other possible virulence molecular markers are shown in **Table 2**. Although the potential virulence mutations are described on the basis of previous studies in animals, the pathogenesis in humans remains unknown. Because the continue evolution of H5N6 viruses in human and poultry (Yuan et al., 2016), patients they had recent history of direct contact with poultry, and drug resistance mutation have been detected in viruses isolated from both human and poultry (Shen et al., 2016), the potential for infection, outbreaks, and pandemic in humans should be closely monitored.

The study had several limitations. Firstly, samples from wild birds were not collected in the study area due to the restriction of regulations. Secondly, environmental samples were not collected to test the presence of avian influenza viruses. Thirdly, the low virus isolation rate could be due to the possible low prevalence of avian influenza viruses and low viral load in samples.

In conclusion, our results suggest transmission of avian influenza viruses from wild birds to poultry. Hence, public health officials must consider methods to disrupt peridomestic novel avian influenza virus transmission in these aquatic bird environments.

AUTHOR CONTRIBUTIONS

M-JM, TZ, Y-HQ, S-HC, G-LW, M-NW, YH, G-YM, L-QF, G-CG, BL, and Y-GT conducted the study and analyzed data, M-JM, S-HC, TZ, G-LW, M-NW, YH, and G-YM processed samples and virus isolation., M-JM and WC conceived and designed the study and wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by the grants from the Program of International Science and Technology Cooperation of China (2013DFA30800), the National Natural Science Foundation of China (81402730), the US National Institute for Allergy and Infection Diseases (R01AI108993-01A1), the Program of Science and Technology of Jiangsu Province (H201448), and the Major Project of Wuxi Health Bureau (Z201404). We thank the farmer who provided the care for sentinel chickens and the several veterinarians who conducted the sampling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01766/full#supplementary-material

Supplementary Figure 1 | Phylogenetic relationships of internal genes of influenza A (H7N2) viruses from sentinel chickens, Wuxi City, Jiangsu Province, China, 2014. Supporting bootstrap values >75 are shown. Red font indicates viruses isolated in present study. PB, polymerase basic; PA, polymerase acidic; NP, nucleoprotein; MP, matrix protein; NS, nonstructural. The black circle dot indicates the viruses isolated in present study.

Supplementary Figure 2 | Phylogenetic relationships of internal genes of influenza A (H5N6) viruses from sentinel chickens, Wuxi City, Jiangsu Province, China, 2014. Supporting bootstrap values >75 are shown. Red font indicates viruses isolated in present study. PB, polymerase basic; PA, polymerase acidic; NP, nucleoprotein; MP, matrix protein; NS, nonstructural. The black circle dot indicates the viruses isolated in present study.

REFERENCES

Bi, Y., Mei, K., Shi, W., Liu, D., Yu, X., Gao, Z., et al. (2015). Two novel reassortants of avian influenza A (H5N6) virus in China. *J. Gen. Virol.* 96, 975–981. doi: 10.1099/vir.0.000056

Cauldwell, A. V., Long, J. S., Moncorgè, O., and Barclay, W. S. (2014). Viral determinants of influenza A virus host range. *J. Gen. Virol.* 95, 1193–1210. doi: 10.1099/vir.0.062836-0

Chinese National Influenza Center (2016). *Chinese Influenza Weekly Report*. National Institute for Viral Disease Control and Prevention,

Chinese Center for Disease Control and Prevention. Available online at: http://www.chinaivdc.cn/cnic/zyzx/lgzb/201606/P020160602468920842502.pdf (Accessed February 1, 2016).

Coman, A., Maftei, D. N., Chereches, R. M., Zavrotchi, E., Bria, P., Dragnea, C., et al. (2014). Avian influenza surveillance in the Danube delta using sentinel geese and ducks. *Influenza Res. Treat.* 2014:965749. doi: 10.1155/2014/965749

Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., et al. (2013). Human infection with a novel avian-origin influenza A (H7N9) virus. *N.Engl. J. Med.* 368, 1888–1897. doi: 10.1056/NEJMoa1304459

Jones, J. C., Sonnberg, S., Kọcer, Z. A., Shanmuganatham, K., Seiler, P., Shu, Y., et al. (2014). Possible role of songbirds and parakeets in transmission of influenza A(H7N9) virus to humans. *Emer. Infect. Dis.* 20, 380–385. doi: 10.3201/eid2003.131271

Jones, J. C., Sonnberg, S., Webby, R. J., and Webster, R. G. (2015). Influenza A(H7N9) virus transmission between finches and poultry. *Emer. Infect. Dis.* 21, 619–628. doi: 10.3201/eid2104.141703

Lam, T. T., Zhou, B., Wang, J., Chai, Y., Shen, Y., Chen, X., et al. (2015). Dissemination, divergence and establishment of H7N9 influenza viruses in China. *Nature* 522, 102–105. doi: 10.1038/nature14348

Li, Z., Chen, H., Jiao, P., Deng, G., Tian, G., Li, Y., et al. (2005). Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* 79, 12058–12064. doi: 10.1128/JVI.79.18.12058-12064.2005

Shen, Y. Y., Ke, C. W., Li, Q., Yuan, R. Y., Xiang, D., Jia, W. X., et al. (2016). Novel reassortant avian influenza A(H5N6) viruses in humans, Guangdong, China, 2015. *Emer. Infect. Dis.* 22, 1507–1509. doi: 10.3201/eid2208.160146

Shi, J., Deng, G., Zeng, X., Kong, H., Wang, X., Lu, K., et al. (2014). Novel influenza A(H7N2) virus in chickens, Jilin province, China, 2014. *Emer. Infect. Dis.* 20, 1719–1722. doi: 10.3201/eid2010.140869

Simon, P., Holder, B. P., Bouhy, X., Abed, Y., Beauchemin, C. A., and Boivin, G. (2011). The I222V neuraminidase mutation has a compensatory role in replication of an oseltamivir-resistant influenza virus A/H3N2 E119V mutant. *J. Clin. Microbiol.* 49, 715–717. doi: 10.1128/JCM.01732-10

Stevens, J., Blixt, O., Tumpey, T. M., Taubenberger, J. K., Paulson, J. C., and Wilson, I. A. (2006). Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312, 404–410. doi: 10.1126/science.1124513

Tobita, K., Sugiura, A., Enomote, C., and Furuyama, M. (1975). Plaque assay and primary isolation of influenza A viruses in an established line of canine kidney cells (MDCK) in the presence of trypsin. *Med. Microbiol. Immunol.* 162, 9–14. doi: 10.1007/BF02123572

World Health Organization (2014). *Real-Time RT-PCR Protocol for the Detection of Avian Influenza A(H7N9) Virus*. Available online at: http://www.who.int/influenza/gisrs_laboratory/cnic_realtime_rt_pcr_protocol_a_h7n9.pdf (Accessed April 13, 2013).

Yu, H., Cowling, B. J., Feng, L., Lau, E. H., Liao, Q., Tsang, T. K., et al. (2013). Human infection with avian influenza A H7N9 virus: an assessment of clinical severity. *Lancet* 382, 138–145. doi: 10.1016/S0140-6736(13)61207-6

Yu, Z., Gao, X., Wang, T., Li, Y., Li, Y., Xu, Y., et al. (2015). Fatal H5N6 avian influenza virus infection in a domestic cat and wild birds in China. *Sci. Rep.* 5:10704. doi: 10.1038/srep10704

Yuan, R., Wang, Z., Kang, Y., Wu, J., Zou, L., Liang, L., et al. (2016). Continuing reassortant of H5N6 subtype highly pathogenic avian influenza virus in Guangdong. *Front. Microbiol.* 7:520. doi: 10.3389/fmicb.2016.00520

Zamarin, D., Ortigoza, M. B., and Palese, P. (2006). Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice. *J. Virol.* 80, 7976–7983. doi: 10.1128/JVI.00415-06

Zhang, H., Li, X., Guo, J., Li, L., Chang, C., Li, Y., et al. (2014). The PB2 E627K mutation contributes to the high polymerase activity and enhanced replication of H7N9 influenza virus. *J. Gen. Virol.* 95, 779–786. doi: 10.1099/vir.0.061721-0

Zhang, Q., Shi, J., Deng, G., Guo, J., Zeng, X., He, X., et al. (2013). H7N9 influenza viruses are transmissible in ferrets by respiratory droplet. *Science* 341, 410–414. doi: 10.1126/science.1240532

Zhao, B., Zhang, X., Zhu, W., Teng, Z., Yu, X., Gao, Y., et al. (2014). Novel avian influenza A(H7N9) virus in tree sparrow, Shanghai, China, 2013. *Emer. Infect. Dis.* 20, 850–853. doi: 10.3201/eid2005.131707

Zhao, K. T., Lu, Y. D., Pu, L. B., and Huang, Z. Y. (1990). A survey on the birds of Suzhou-Wuxi area, Jiangsu province. *J. Suzhou Railway Teach. Coll.* 7, 46–61.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhao, Qian, Chen, Wang, Wu, Huang, Ma, Fang, Gray, Lu, Tong, Ma and Cao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 April 2017
doi: 10.3389/fmicb.2017.00584

Molecular Dynamics Simulation of the Influenza A(H3N2) Hemagglutinin Trimer Reveals the Structural Basis for Adaptive Evolution of the Recent Epidemic Clade 3C.2a

*Masaru Yokoyama[1], Seiichiro Fujisaki[2], Masayuki Shirakura[2], Shinji Watanabe[2], Takato Odagiri[2], Kimito Ito[3] and Hironori Sato[1]**

[1] Laboratory of Viral Genomics, Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan, [2] Influenza Virus Research Center, National Institute of Infectious Diseases, Tokyo, Japan, [3] Research Center for Zoonosis Control, Hokkaido University, Hokkaido, Japan

OPEN ACCESS

Edited by:
Aeron Hurt, WHO Collaborating Centre for Reference and Research on Influenza, Australia

Reviewed by:
Hirotaka Ode, National Hospital Organization Nagoya Medical Center, Japan
Hidekatsu Iha, Oita University, Japan

***Correspondence:**
Hironori Sato
hirosato@nih.go.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *05 January 2017*
Accepted: *21 March 2017*
Published: *10 April 2017*

Citation:
Yokoyama M, Fujisaki S, Shirakura M, Watanabe S, Odagiri T, Ito K and Sato H (2017) Molecular Dynamics Simulation of the Influenza A(H3N2) Hemagglutinin Trimer Reveals the Structural Basis for Adaptive Evolution of the Recent Epidemic Clade 3C.2a. Front. Microbiol. 8:584. doi: 10.3389/fmicb.2017.00584

Influenza A(H3N2) has been a major cause of seasonal influenza in humans since 1968, and has evolved by antigenic drift under the constantly changing human herd immunity. Increasing evidence suggests that the antigenic change occasionally occurred concomitant with the alterations of the N-glycosylation site profile and hemagglutination activity of the virion surface protein hemagglutinin (HA). However, the structural basis of these changes remains largely unclear. To address this issue, we performed molecular dynamics simulations of the glycosylated HA trimers of the A(H3N2), which has a novel pattern of Asn-X-Ser/Thr sequons unique in the new A(H3N2) epidemic clade 3C.2a and is characterized by attenuated ability to agglutinate nonhuman erythrocytes. Comparison of the equilibrated structures of the glycosylated HA trimers with and without the 3C.2a-specific mutations reveals that the mutations could induce a drastic reduction in the apical space for the ligand binding via glycan-shield rearrangement. The results suggest that the 3C.2a strain has evolved an HA structure that is advantageous for evading pre-existing antibodies, while also increasing the ligand binding specificity. These findings have structural implications for our understanding of the phenotypic changes, evolution, and fate of influenza A(H3N2).

Keywords: MD simulation, influenza A(H3N2), HA protein, N-linked glycans, mutations, structural change

INTRODUCTION

The hemagglutinin (HA) protein of influenza virus is a glycosylated type I integral membrane protein that protrudes from the mature virion surface and plays critical roles in viral interactions with hosts. The HA protein is synthesized in infected cells as a precursor HA0, and is subsequently cleaved by cellular proteases to HA1 and HA2 subunits that are covalently attached by a disulfide bond. The mature HA protein on the virion is composed of three pairs of the HA1/HA2 subunits (Ha et al., 2003). The tip of the HA protein forms a globular structure, termed the globular head, and confers on the virus an ability to attach cells via interactions with the sialic acid-containing

glycan moiety on the target cell surface (Ha et al., 2003). Meanwhile, the HA globular head constitutes the major viral antigenic sites that induce neutralization antibodies in infected hosts. These functional and antigenic features drive sequence and structural variations, particularly near the receptor-binding site in the globular head, according to specific rules (Smith et al., 2004; Koel et al., 2013). Importantly, the sequence variation on the globular head causes various phenotypic changes of viruses, including changes in antigenicity and receptor specificity. Therefore, it is critical to determine the structural changes in the HA globular head in order to understand the viral interplay with the hosts and evolution. Unfortunately, however, it is usually time consuming to characterize mutation-induced structural changes by experimental approaches alone.

Computational science is a rapidly growing area that now successfully complements the experimental and theoretical sciences in various fields, including life science. For example, recent advances in molecular dynamics (MD) simulation enable us to characterize changes in the three-dimensional structures of the mutated proteins in relatively short timescales compared with the experimental approaches (Ode et al., 2012; Sato et al., 2013). The MD simulations have been used to disclose the structural basis of the adaptation and evolution of the highly mutable human immunodeficiency virus (HIV). This includes elucidation of the HIV structural changes associated with the phenotypic changes in viral neutralization sensitivity and receptor tropism (Naganawa et al., 2008; Yokoyama et al., 2012, 2016; Kuwata et al., 2013), viral sensitivity to antiviral protein (Miyamoto et al., 2012), viral drug sensitivity (Yuan et al., 2013), viral growth in non-natural host cells (Yokoyama et al., 2016), and viral sensitivities to antibodies by drug-resistance mutations (Alam et al., 2016; Hikichi et al., 2016).

In this study, we used the MD simulation to gain new insights into the roles of mutations in a recent epidemic variant of the influenza A(H3N2) viruses. The A(H3N2) viruses have emerged on 1968 in humans of southern Asia and were soon widespread in the world. Thereafter, the A(H3N2) has been a major cause of seasonal influenza in humans to date. During the 2014/15 epidemic season of influenza, a new A(H3N2) substrain had rapidly predominated in humans worldwide (Skowronski et al., 2016). Notably, the hemagglutination activity of this substrain somehow could be measured with only a small portion of the viral population using a conventional hemagglutination assay with nonhuman erythrocytes (Skowronski et al., 2016). The A(H3N2) substrain is characterized by alterations of the N-glycosylation sequons on the globular heads of the HA protein as compared with other A(H3N2) clades (Skowronski et al., 2016) and is now referred to as 3C.2a. The oligosaccharides on the HA protein play key roles in viral antigenicity (Aytay and Schulze, 1991; Abe et al., 2004; Saito et al., 2004; Ping et al., 2008; Das et al., 2010; Wang et al., 2010; Wanzeck et al., 2011) and binding specificity/affinity to the cellular receptor (Gunther et al., 1993; Ohuchi et al., 1997; Gambaryan et al., 1998; Matrosovich et al., 1999; Tsuchiya et al., 2002; Wang et al., 2009; de Vries et al., 2010; Liao et al., 2010). However, it remains unclear how the 3C.2a mutations altered the HA structure and attenuated the hemagglutination activity with nonhuman erythrocytes. To address this issue, we here examined the structural effects of the four mutations in the globular heads using MD simulations. The obtained results predicted that the mutations could induce rearrangement of the glycan shield around the receptor-binding surface of the HA protein, leading to shrinkage of the ligand-accessible space.

MATERIALS AND METHODS

Genetic Clades Determination

Genetic clades determination of the influenza virus in Japan has been performed routinely as the part of the work of the National Epidemiological Surveillance of Infectious Diseases in Japan[1] and the Global Surveillance of Influenza in the WHO Reference Laboratories[2]. Information on the specimen collection and weekly report of HA type is available on the National Epidemiological Surveillance of Infectious Diseases[3]. Briefly, RNAs were extracted from viruses by using QIAamp viral RNA kit (QIAGEN, Dusseldorf, Germany). The HA genes were amplified from extracted RNAs by RT-PCRs using gene-specific primers (the primer sequences are available upon request) and SuperScript III One-step RT-PCR system with Platinum Taq (Invitrogen, Carlsbad, CA, USA). Sequencing reactions were performed with BigDye terminator kit (Applied Biosystems) and sequences were determined using 3730xl DNA analyzer (Applied Biosystems, Foster City, CA, USA). The genetic clades were determined by the generation of phylogenetic trees of the HA genes. The phylogenetic trees were constructed using MEGA 6 software (Tamura et al., 2013) with the neighbor-joining method. The numbers of HA sequences obtained in the individual periods are 39, 83, 125, 108, 70, and 77 for the September 2013 – Jan. 2014, February 2014 – August 2014, September 2014 – January 2015, February 2015 – August 2015, September 2015 – January 2016, and February 2016 – August 2016, respectively. The nucleotide sequences used in this study are registered at GISAID, a publicly accessible influenza virus database[4]. Accession number at GISAID for the HA sequence used for the molecular modeling in this study is EPI543763 (A/Switzerland/9715293/2013 strain).

Molecular Modeling of a Glycosylated HA Trimer in the Ligand-free State

Three-dimensional (3-D) models for glycosylated extracellular domains of HA trimers of influenza A (H3N2) in the ligand-free state were constructed by homology modeling with Molecular Operating Environment (MOE) (Chemical Computing Group Inc., Montreal, QC, Canada). The crystal structure of the HA trimer of the influenza A/Victoria/361/2011 (H3N2) virus (PDB code: 4O5N; resolution: 1.75 Å; amino acid residues 4–325 and 330–502 for HA1 and HA2 peptides, respectively) was used as

[1]http://www.nih.go.jp/niid/en/influenza-e.html

[2]http://www.who.int/influenza/en/

[3]http://www.nih.go.jp/niid/en/influenza-e/2099-idsc/iasr-flu-e/6791-iasr-inf-e20160925.html

[4]http://platform.gisaid.org/epi3/frontend

the modeling template. Obtained models were optimized by energy minimization using MOE and an Amber10: Extended Huckel Theory (EHT) force field implemented in MOE, which combines Amber10 and EHT bonded parameters for the large-scale energy minimization (Gerber and Muller, 1995; Case et al., 2005). The high-mannose oligosaccharide $Man_5GlcNAc_2$ was added to potential *N*-glycosylation sites in HA using Online Glycoprotein Builder[5].

MD Simulation of Glycosylated HA Trimer Models

Glycosylated HA trimer models in a ligand-free state were subjected to MD simulation essentially as described for simulations of HIV-1 gp120 (Yokoyama et al., 2016). MD simulations were performed by the PMEMD (Particle Mesh Ewald Molecular Dynamics) module in the AMBER 14 program package (Case et al., 2014), employing the Amber ff99SB-ILDN force field, a protein force field with improved side-chain torsion potentials (Lindorff-Larsen et al., 2010), the GLYCAM06 force field, a biomolecular force field for glycans (Kirschner et al., 2008), and the TIP3P water model for simulations of aqueous solutions (Jorgensen et al., 1983). Bond lengths involving hydrogen were constrained with SHAKE, a constraint algorithm to satisfy a Newtonian motion (Ryckaert et al., 1977), and the time step for all MD simulations was set to 2 fs. A non-bonded cutoff of 10 Å was used. After heating calculations for 20 ps until 310 K using the NVT ensemble for the constant volume, temperature, and numbers of particles in the system, simulations were executed using the NPT ensemble for the constant pressure, temperature, and numbers of particles in the system at 1 atm, at 310 K, and in 150 mM NaCl for 100 ns. Root mean square deviations (RMSDs) between the heavy atoms of the two superposed proteins were used to measure the overall structural differences between the two proteins (Case et al., 2005). The RMSD was calculated using the cpptraj module in AmberTools 14, a trajectory analysis tool (Case et al., 2014). We used "Computer System for the Prediction of Mutations of Pathogens" at Research Center for Zoonosis Control, Hokkaido University for the MD simulations.

Calculation of Root Mean Square Fluctuation (RMSF)

We calculated RMSFs of individual components of the high mannose oligosaccharides $Man_5GlcNAc_2$ around the receptor binding site between 50 to 100 ns of MD simulations to quantify structural dynamics of glycans during the MD simulations. The average structures during the last 50 ns of MD simulations were used as reference structures for RMSF calculation. RMSFs were calculated as previously described (Naganawa et al., 2008; Yokoyama et al., 2012, 2016; Kuwata et al., 2013) by using the ptraj module in Amber, a trajectory analysis tool (Case et al., 2005).

[5]http://glycam.org/

RESULTS

Temporal Dynamics of Clade Populations of A(H3N2) Viruses Since September 2013

Over the 2014/15 influenza season, influenza surveillance reports from reference laboratories warned of a global epidemic of a new A(H3N2) variant population, termed 3C.2a. In Japan, the 3C.2a was first detected as a relatively minor population over the season from February to August in 2014, during which four A(H3N2) clades co-existed (**Figure 1**, upper panel). However, the 3C.2a rapidly became dominant at the beginning of the winter season in 2015, displaced pre-existing clades, and has continually predominated at the collection sites in Japan ever since, representing 76 of 77 (98.7%) clades in the season from February to August in 2016 (**Figure 1**, lower panel). In parallel, another A (H3N2) clade, 3C.3a, which co-existed in the same period as the 3C. 2a over the February to August 2014 season, became minor after February 2015 (**Figure 1**). These results are consistent with the reports on the global epidemic of the 3C.2a from the WHO Reference Laboratories[6], and suggest the selective advantage of the 3C.2a for human-to-human transmission during the study period. Notably, hemagglutination activity of the 3C.2a was significantly attenuated when measured with a conventional hemagglutination assay (Skowronski et al., 2016). Together, these data suggest that certain structural changes occurred in the HA protein to confer selective advantages on the 3C.2a.

Amino Acid Signatures of the 3C.2a HA Protein

Molecular epidemiological data from above study and the global surveillance of influenza in the WHO Reference Laboratories suggest that the A (H3N2) clades, 3C.2a and 3C.3a, were diversified from 3C.2 and 3C.3, respectively and thereafter the 3C.2a had dominated over the 3C.3a. Therefore, we examined differences in HA proteins between the clades 3C.2a and 3C.3a. The HA protein of the 3C.2a population has multiple amino acid substitutions as compared with the clade 3C.3a (amino acid numbers 3, 128, 138, 142, 144, 159, 160, 311, 326, and 489). In this study, we focused on the four substitutions around the receptor-binding surface on the tip of the HA protein, i.e., Ala128Thr, Asn144Ser, Ser159Tyr, and Lys160Thr (**Figure 2**). The Ala128Thr and Lys160Thr create new potential N-glycosylation sites, Asn-X-Ser/Thr, whereas Asn144Ser results in the loss of a single N-glycosylation site. Ser159Tyr is located adjacent to the Lys160Thr. The Ala128Thr substitution initially detected in the 3C.2 and had been preserved in its descendent, 3C.2a, whereas the other three substitutions newly emerged in the 3C.2a (**Figure 2A**). All the mutations are placed at or near the major antigenic elements for H3 HAs, such as antigen sites A (amino acids 140-146) and B (amino acids 155-160 and 188-198) (Webster and Laver, 1980; Gerhard et al., 1981; Wiley et al., 1981; Knossow et al., 1984; Wiley and Skehel, 1987) (**Figure 2B**).

[6]http://www.who.int/influenza/en/

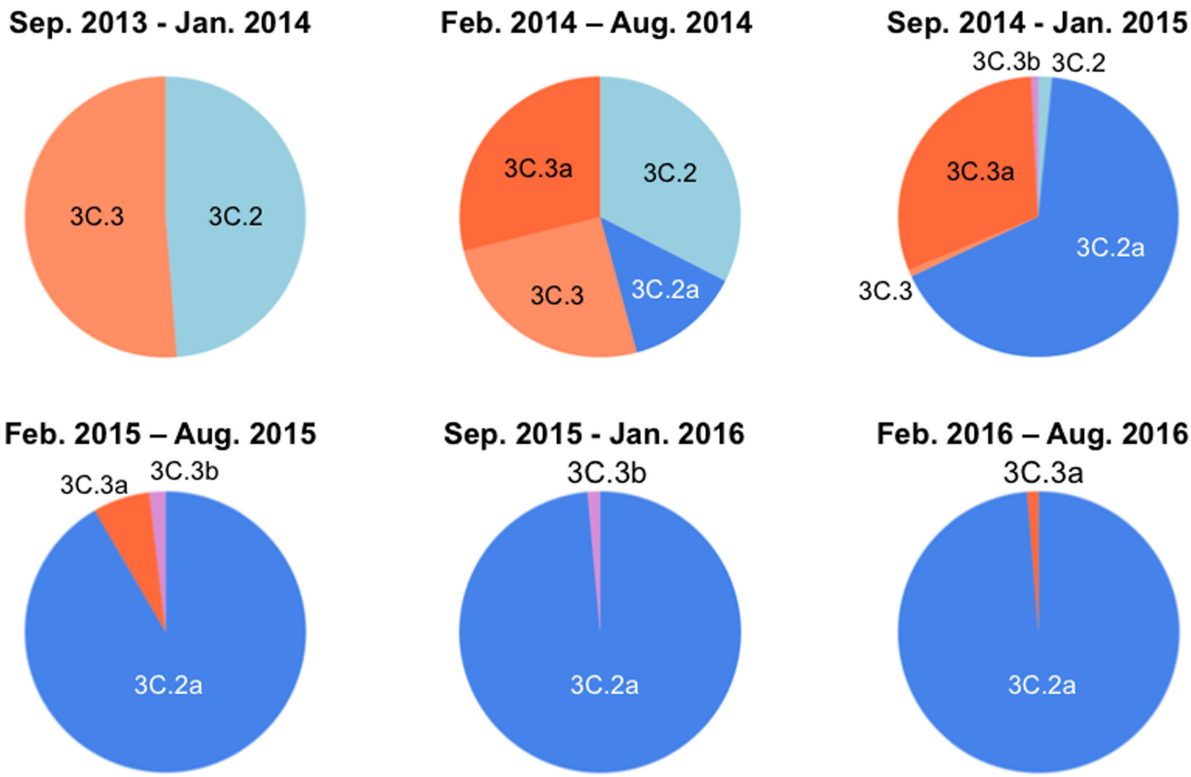


FIGURE 1 | Temporal dynamics of clade populations of A (H3N2) viruses since September 2013. The clade-populations based on the phylogenetic tree of the hemagglutinin (HA) genes are shown by pie charts. The periods of each pie chart are determined by the collection dates of specimens. The clades are colored as follows: 3C.2, light blue; 3C.2a, blue; 3C.3, light orange; 3C.3a, orange; 3C.3b, light magenta.

MD Simulations of Glycosylated HA Trimers in the Ligand-free State

To understand the structural impacts of the four mutations on the tip of the 3C.2a HA protein, we constructed molecular models of the glycosylated HA trimers with and without the mutations. Two structural models of the glycosylated HA trimers in a ligand-free state were constructed by homology modeling: a model for the A(H3N2) 2015/16 vaccine strain (A/Switzerland/9715293/2013) that belongs to the perishing clade 3C.3a (**Figures 1**, **2**) and a model for A/Switzerland/9715293/2013 possessing the four 3C.2a-specific mutations on the HA globular heads (Ala128Thr, Asn144Ser, Ser159Tyr, and Lys160Thr). The obtained models were optimized by energy minimization using the MOE:EHT force field (Gerber and Muller, 1995; Case et al., 2005) and were subjected to the MD simulations using the Amber ff99SB-ILDN force field (Lindorff-Larsen et al., 2010) and the GLYCAM06 force field (Kirschner et al., 2008) as described previously (Yokoyama et al., 2016).

The structural dynamics of the glycosylated HA proteins in solution were monitored by RMSD between the initial model structure and the structures at given time points of the MD simulation (**Figure 3**). For the protein portions of the glycosylated HA molecules, the RMSDs sharply increased in the beginning and reached a near plateau after 20 ns of the MD simulations in both the 3C.3a and its mutant with 3C.2a mutations (**Figures 3A,B**). This RMSD profile was similar to that obtained with the glycosylated HIV-1 gp120 protein (Yokoyama et al., 2016). In addition, we monitored the RMSDs of the glycan portions of the glycosylated HA molecules. We obtained an RMSD profile that was similar to that of the protein portions: there was a sharp increase soon after the MD simulation onset followed by a near plateau after 20 ns (**Figures 3C,D**). These results suggest that the structural distortions of the amino acid residues and glycans of the initial models were relieved shortly after the start of MD simulation under thermodynamic driving forces in solution. The data also predict that the HA structure can reach a state of thermodynamic equilibrium in solution.

Finally, we compared the 3-D structure of the HA trimer in a state of thermodynamic equilibrium during the MD simulations (100 ns) between the 3C.3a and its mutant (**Figure 4A**). A marked structural difference between the 3C.3a and its mutants was detected on the globular heads of the HA protein at 100 ns of two MD simulations: the glycan moieties were placed more compactly in the mutant than in the 3C.3a (**Figures 4A,B**). Consequently, the space for the access of ligands around the receptor-binding site was more confined in the mutant (**Figure 4C**). The apical spaces for ligand binding could be influenced not only by the arrangements of glycans but also by fluctuation of the glycans. Therefore, we examined structural fluctuations of the

A

Clade	Positions and types of amino acid substitutions for N-glycosylation sequon changes			
	128	144	159	160
3C.3	Ala	Asn	Phe	Lys
3C.3a	Ala	Asn	Ser	Lys
3C.3b	Ala	Asn	Phe	Lys
3C.2	Thr	Asn	Phe	Lys
3C.2a	Thr	Ser	Tyr	Thr

B

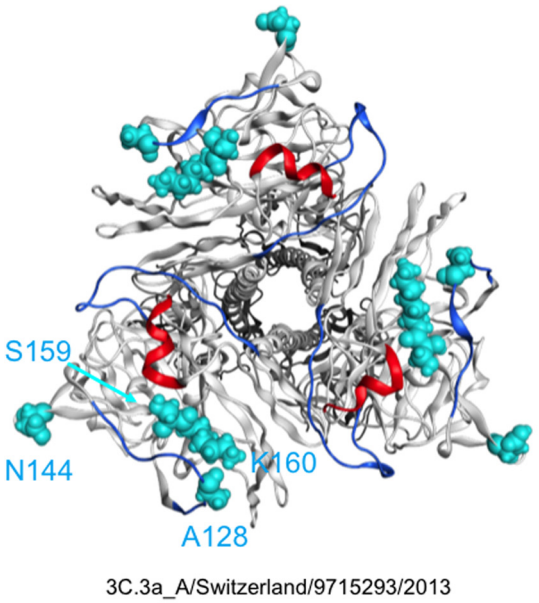


FIGURE 2 | Amino acid signatures of the 3C.2a HA protein. Types and 3-D locations of amino acid substitutions around the receptor-binding surface on the tip of the HA protein are shown. **(A)** Types of amino acid residues among five A(H3N2) variant strains emerged in the 2013–2016 periods in Japan. **(B)** Locations of different amino acid residues between 3C.3a and 3C.2a are highlighted by cyan color on the HA protein model of the 3C.3a strain (A(H3N2) 2015/16 vaccine strain: A/Switzerland/9715293/2013).

individual glycans around the receptor binding site by calculating RMSF using snapshots of structures from 50 to 100 ns of each MD simulation. In the present HA models, the high mannose oligosaccharides $Man_5GlcNAc_2$ were attached on the asparagine residues of potential *N*-glycosylation sites around the receptor binding site (**Figure 5A**). We calculated RMSFs of individual components of the oligosaccharides, such as GlcNAc and Man, at positions 1 to 7 (**Figure 5B**). Notably, RMSFs of the glycan components at given positions, and the profiles of RMSFs at the seven positions were similar among the four *N*-glycosylation sites. The data indicate that all glycans fluctuated in a similar fashion around the receptor binding site during the MD simulations. These results suggest that the apical spaces for ligand binding are mainly influenced by arrangements of glycans and that the spaces are constantly different between the two models during the MD simulations.

DISCUSSION

In this report, we studied the structural impact of mutations residing in the HA globular head of a recent epidemic clade population of the A(H3N2). This clade, termed 3C.2a, predominated over other co-existing A(H3N2) clades including the clade 3C.3a during the 2014/15 and 2015/16 seasons, and retained a unique set of sequon mutations in the HA globular heads (**Figures 1**, **2**). To analyze the HA structures under near physiological conditions, we performed MD simulations of the glycosylated HA trimers and obtained structures in a state of thermodynamic equilibrium in solution (**Figure 3**). Comparisons of the obtained structures revealed marked changes in the glycan shield around the receptor-binding site (**Figures 4**, **5**). This finding has important implications for our understanding of phenotypic changes, evolution, and fate of the influenza virus A(H3N2).

First, the present results indicate that the 3C.2a-specific mutations have an impact on the immunological features of the HA protein. Our MD simulations show that the glycans on the HA globular head are rearranged by the mutations into a configuration markedly distinct from that of the 2015/16 vaccine strain (**Figures 4**, **5**). Therefore, these mutations could induce changes in the HA antigenicity. These structural findings are consistent with previous reports demonstrating critical roles of the oligosaccharides on the HA protein in the viral antigenicity of influenza viruses (Aytay and Schulze, 1991; Abe et al., 2004; Saito et al., 2004; Ping et al., 2008; Das et al., 2010; Wang et al., 2010; Wanzeck et al., 2011). Moreover, we found that the glycan rearrangement resulted in shrinkage of the access space on the top of the HA protein near the receptor-binding site of the globular head (**Figure 4**). This structural change could cause steric hindrance for binding of the antibodies directed to the antigen site B. Thus, it is likely that the HA protein of the 3C.2a had a selective advantage in evading antibodies against receptor-binding site as compared with the 3C.3a HA. These possibilities are consistent with the rapid spread and predominance of the clade C3.2a over the clade 3C.3a in human populations during the study period (**Figure 1**) (Skowronski et al., 2016). Antigenic analysis of these viruses should be done to make sure these possibilities.

Secondly, the present results indicate that the 3C.2a-specific mutations impact the receptor specificity of the HA protein. Our MD simulations showed that the new arrangement of glycans on the HA globular head could shrink the space for the access of the sialic acid-containing glycan moiety on the target cell surface (**Figure 4**). This structural change could cause an increase in the ligand specificity and/or affinity of the HA protein, and thereby lead to a preference for receptors on the human respiratory organs, but not for receptors on the nonhuman erythrocytes. These structural findings are consistent with the attenuation of hemagglutination activity of the 3C.2a HA protein when assessed with a conventional hemagglutination assay using nonhuman erythrocytes (Skowronski et al., 2016). Similarly, previous studies have demonstrated that the oligosaccharides on the HA protein surface play key roles in the binding specificity and the affinity

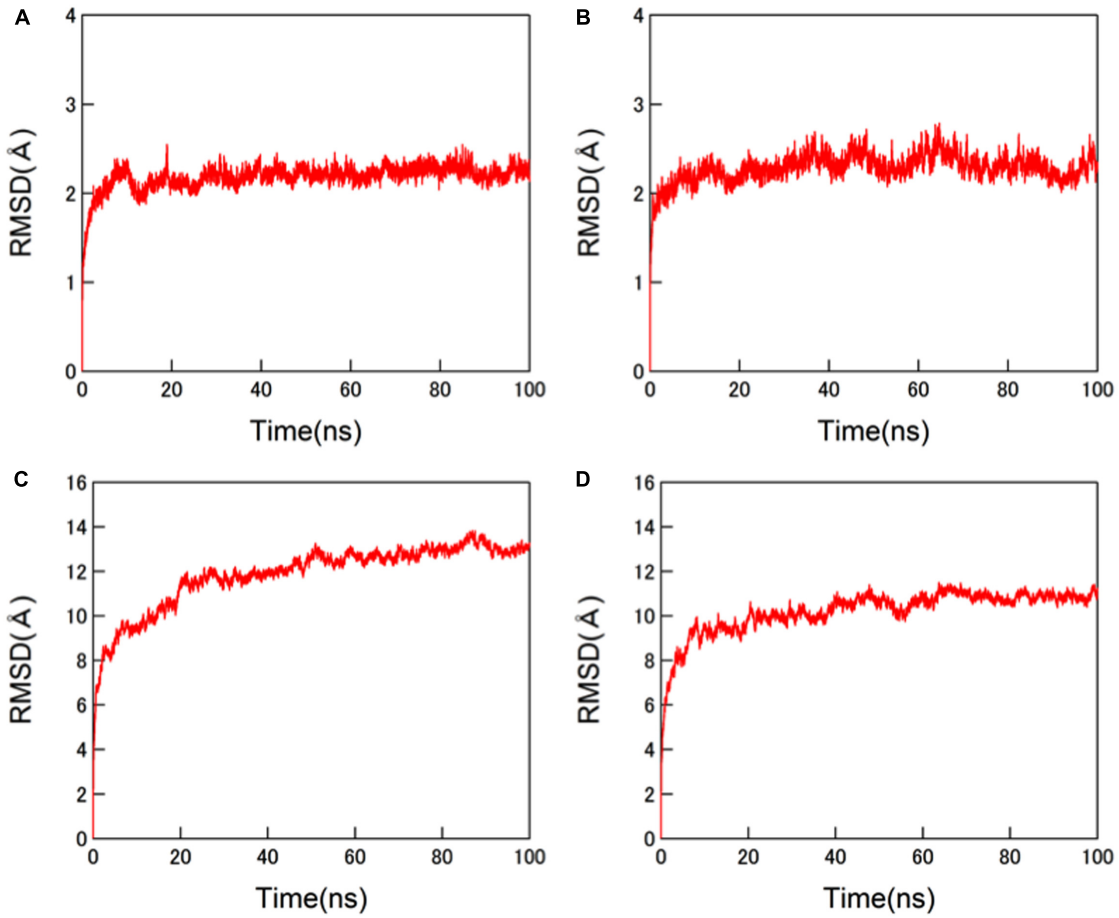


FIGURE 3 | Molecular dynamics (MD) simulations of glycosylated HA trimers in a ligand-free state. Molecular models of glycosylated extracellular domains of HA trimer in a ligand-free state were constructed by homology modeling using MOE (Chemical Computing Group Inc., Montreal, QC, Canada) and subjected to MD simulations using the AMBER 14 program package (Case et al., 2014) as described in the Section "Materials and Methods". The structural dynamics of the HA trimers in solution were monitored by root mean square deviations (RMSD) between the initial model structure and the structures at given time points of MD simulation. **(A,C)** HA trimer of a representative strain (A/Switzerland/9715293/2013) belonging to the subclade 3C.3a. **(B,D)** HA trimer of the A/Switzerland/9715293/2013 with Ala128Thr, Asn144Ser, Ser159Tyr, and Lys160Thr substitutions. **(A,B)** RMSDs of the protein portions. **(C,D)** RMSDs of N-glycans.

to infection receptors (Gunther et al., 1993; Ohuchi et al., 1997; Gambaryan et al., 1998; Matrosovich et al., 1999; Tsuchiya et al., 2002; Wang et al., 2009; de Vries et al., 2010; Liao et al., 2010).

Thirdly, the present findings have implications in terms of the adaptive evolution of A (H3N2). Previous reports have highlighted the importance of changes in the glycan shield for viral adaptation (Deom et al., 1986). Interestingly, the numbers of potential N-glycosylation sites in the HA protein have been continuously increasing since 1968: only two sequons existed –in the initial strains, whereas more than 7–10 sequons are common in the present epidemic strains (Blackburne et al., 2008). This change is likely to be a basic strategy by which A (H3N2) has maintained its presence in human populations with changing herd immunity over the last 48 years. However, the change also runs the risk of creating an evolutionary "dead-end" for the virus, because the acquisition of new glycans on the HA globular head increases the chances for steric hindrance during receptor binding. Notably, both the present and recent studies suggest that the HA protein possessing the 3C.2a-type glycan shield would significantly affect ligand specificity (**Figure 4**) (Skowronski et al., 2016). Thus, the 3C.2a has evolved an HA structure that is advantageous for evading pre-existing antibodies, while also reducing ligand affinity to nonhuman glycan moieties.

Finally, the present have structural implications in the fate of influenza A (H3N2) viruses. Our study indicates that an increase in the number of the glycan moieties around the receptor-binding site could induce a drastic reduction in the apical space for ligand binding (**Figure 4**). As discussed above, continuous interactions between A (H3N2) and human immunity seem to force the virus to serially increase the numbers of glycans on the HA protein. This eventually may reduce the HA binding affinity via an increase in steric hindrance even to the human ligands, and thereby decrease the replication fitness of the virus for spread among humans. Thus, if the overall tendency of HA to increase the numbers of glycosylation sites (Blackburne et al., 2008) continues in the future, it may

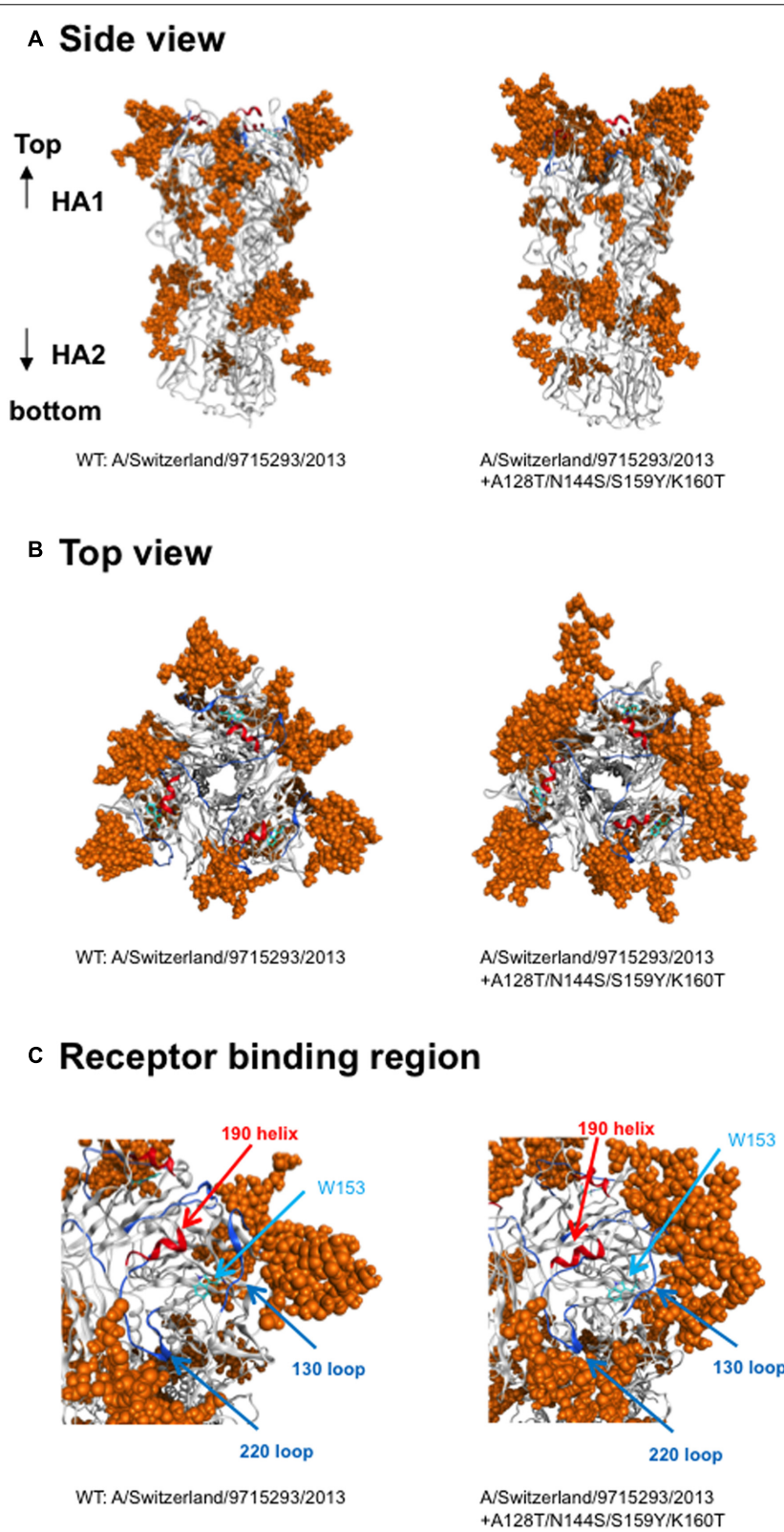


FIGURE 4 | Comparison of HA trimer models at 100 ns of MD simulations. The 3-D structure of the HA trimer in a state of thermodynamic equilibrium during the MD simulations (100 ns) was compared between the vaccine strain and its mutant. **(A)** Side view, **(B)** top view. **(C)** Receptor-binding region. Helix and loops in the receptor-binding site are marked by red and blue colors, respectively. The tryptophan at position 153 for the binding of sialic acid is shown.

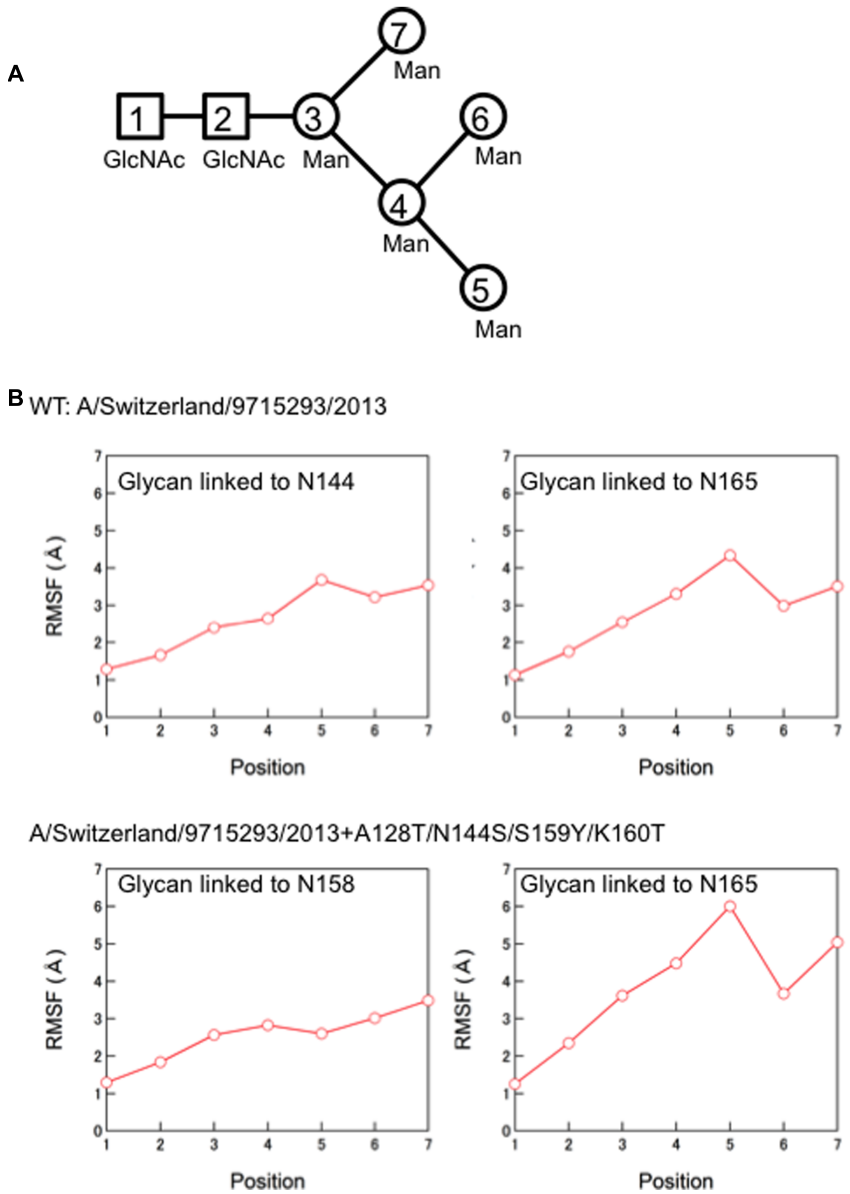


FIGURE 5 | Molecular dynamics of the individual glycans around the receptor binding site. (A) Illustration of components of the high mannose oligosaccharides $Man_5GlcNAc_2$. **(B)** Root Mean Square Fluctuation (RMSF) of oligosaccharides. Structural fluctuations of the individual components of glycans around the receptor binding site of the HA proteins were analyzed by calculating RMSF using snapshots of structures from 50 to 100 ns of each MD simulation. The average structures during the last 50 ns of MD simulations were used as reference structures for RMSF calculation.

eventually put the influenza A (H3N2) viruses in an evolutionary cul-de-sac.

AUTHOR CONTRIBUTIONS

MY, SW, TO, and HS conceived and designed the study. SF and MS performed sequencing and phylogenetic analysis. KI and HS prepared the computing environment. MY performed MD simulations. HS prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by a Grant-in-Aid for emerging and reemerging infectious diseases to HS (Grant Number: H27-Shinkogyousei-shitei-002) from the Ministry of Health, Labour and Welfare of Japan, a financial support from the Joint Usage/Research Center Program at Research Center for Zoonosis Control, Hokkaido University, a Grant-in-Aid for Scientific Research (C) to MY (Grant Number:16K08824) from the Ministry of Education, Culture, Sports, Science and Technology, and a Grant-in-Aid for the Research Program

on Re-emerging Infectious Diseases to TO from the Japan Agency for Medical Research and development, AMED.

ACKNOWLEDGMENT

We thank Ms. Hiromi Nakamura for technical assistance.

REFERENCES

Abe, Y., Takashita, E., Sugawara, K., Matsuzaki, Y., Muraki, Y., and Hongo, S. (2004). Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *J. Virol.* 78, 9605–9611. doi: 10.1128/JVI.78.18.9605-9611.2004

Alam, M., Kuwata, T., Shimura, K., Yokoyama, M., Ramirez Valdez, K. P., Tanaka, K., et al. (2016). Enhanced antibody-mediated neutralization of HIV-1 variants that are resistant to fusion inhibitors. *Retrovirology* 13:70. doi: 10.1186/s12977-016-0304-7

Aytay, S., and Schulze, I. T. (1991). Single amino acid substitutions in the hemagglutinin can alter the host range and receptor binding properties of H1 strains of influenza A virus. *J. Virol.* 65, 3022–3028.

Blackburne, B. P., Hay, A. J., and Goldstein, R. A. (2008). Changing selective pressure during antigenic changes in human influenza H3. *PLoS Pathog.* 4:e1000058. doi: 10.1371/journal.ppat.1000058

Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., et al. (2014). *AMBER 14*. San Francisco, CA: University of California.

Case, D. A., Cheatham, T. E. III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., et al. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688. doi: 10.1002/jcc.20290

Das, S. R., Puigbo, P., Hensley, S. E., Hurt, D. E., Bennink, J. R., and Yewdell, J. W. (2010). Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain. *PLoS Pathog.* 6:e1001211. doi: 10.1371/journal.ppat.1001211

de Vries, R. P., de Vries, E., Bosch, B. J., de Groot, R. J., Rottier, P. J., and de Haan, C. A. (2010). The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity. *Virology* 403, 17–25. doi: 10.1016/j.virol.2010.03.047

Deom, C. M., Caton, A. J., and Schulze, I. T. (1986). Host cell-mediated selection of a mutant influenza A virus that has lost a complex oligosaccharide from the tip of the hemagglutinin. *Proc. Natl. Acad. Sci. U.S.A.* 83, 3771–3775. doi: 10.1073/pnas.83.11.3771

Gambaryan, A. S., Marinina, V. P., Tuzikov, A. B., Bovin, N. V., Rudneva, I. A., Sinitsyn, B. V., et al. (1998). Effects of host-dependent glycosylation of hemagglutinin on receptor-binding properties on H1N1 human influenza A virus grown in MDCK cells and in embryonated eggs. *Virology* 247, 170–177. doi: 10.1006/viro.1998.9224

Gerber, P. R., and Muller, K. (1995). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des.* 9, 251–268. doi: 10.1007/BF00124456

Gerhard, W., Yewdell, J., Frankel, M. E., and Webster, R. (1981). Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. *Nature* 290, 713–717. doi: 10.1038/290713a0

Gunther, I., Glatthaar, B., Doller, G., and Garten, W. (1993). A H1 hemagglutinin of a human influenza A virus with a carbohydrate-modulated receptor binding site and an unusual cleavage site. *Virus Res.* 27, 147–160. doi: 10.1016/0168-1702(93)90078-2

Ha, Y., Stevens, D. J., Skehel, J. J., and Wiley, D. C. (2003). X-ray structure of the hemagglutinin of a potential H3 avian progenitor of the 1968 Hong Kong pandemic influenza virus. *Virology* 309, 209–218. doi: 10.1016/S0042-6822(03)00068-0

Hikichi, Y., Yokoyama, M., Takemura, T., Fujino, M., Kumakura, S., Maeda, Y., et al. (2016). Increased HIV-1 sensitivity to neutralizing antibodies by mutations in the Env V3-coding region for resistance to CXCR4 antagonists. *J. Gen. Virol.* 97, 2427–2440. doi: 10.1099/jgv.0.000536

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., et al. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* 29, 622–655. doi: 10.1002/jcc.20820

Knossow, M., Daniels, R. S., Douglas, A. R., Skehel, J. J., and Wiley, D. C. (1984). Three-dimensional structure of an antigenic mutant of the influenza virus haemagglutinin. *Nature* 311, 678–680. doi: 10.1038/311678a0

Koel, B. F., Burke, D. F., Bestebroer, T. M., van der Vliet, S., Zondag, G. C., Vervaet, G., et al. (2013). Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* 342, 976–979. doi: 10.1126/science.1244730

Kuwata, T., Takaki, K., Yoshimura, K., Enomoto, I., Wu, F., Ourmanov, I., et al. (2013). Conformational epitope consisting of the V3 and V4 loops as a target for potent and broad neutralization of simian immunodeficiency viruses. *J. Virol.* 87, 5424–5436. doi: 10.1128/JVI.00201-13

Liao, H. Y., Hsu, C. H., Wang, S. C., Liang, C. H., Yen, H. Y., Su, C. Y., et al. (2010). Differential receptor binding affinities of influenza hemagglutinins on glycan arrays. *J. Am. Chem. Soc.* 132, 14849–14856. doi: 10.1021/ja104657b

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78, 1950–1958. doi: 10.1002/prot.22711

Matrosovich, M., Zhou, N., Kawaoka, Y., and Webster, R. (1999). The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. *J. Virol.* 73, 1146–1155.

Miyamoto, T., Nakayama, E. E., Yokoyama, M., Ibe, S., Takehara, S., Kono, K., et al. (2012). The carboxyl-terminus of human immunodeficiency virus type 2 circulating recombinant form 01_AB capsid protein affects sensitivity to human TRIM5alpha. *PLoS ONE* 7:e47757. doi: 10.1371/journal.pone.0047757

Naganawa, S., Yokoyama, M., Shiino, T., Suzuki, T., Ishigatsubo, Y., Ueda, A., et al. (2008). Net positive charge of HIV-1 CRF01_AE V3 sequence regulates viral sensitivity to humoral immunity. *PLoS ONE* 3:e3206. doi: 10.1371/journal.pone.0003206

Ode, H., Nakashima, M., Kitamura, S., Sugiura, W., and Sato, H. (2012). Molecular dynamics simulation in virus research. *Front. Microbiol.* 3:258. doi: 10.3389/fmicb.2012.00258

Ohuchi, M., Ohuchi, R., Feldmann, A., and Klenk, H. D. (1997). Regulation of receptor binding affinity of influenza virus hemagglutinin by its carbohydrate moiety. *J. Virol.* 71, 8377–8384.

Ping, J., Li, C., Deng, G., Jiang, Y., Tian, G., Zhang, S., et al. (2008). Single-amino-acid mutation in the HA alters the recognition of H9N2 influenza virus by a monoclonal antibody. *Biochem. Biophys. Res. Commun.* 371, 168–171. doi: 10.1016/j.bbrc.2008.04.045

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341. doi: 10.1016/0021-9991(77)90098-5

Saito, T., Nakaya, Y., Suzuki, T., Ito, R., Saito, T., Saito, H., et al. (2004). Antigenic alteration of influenza B virus associated with loss of a glycosylation site due to host-cell adaptation. *J. Med. Virol.* 74, 336–343. doi: 10.1002/jmv.20178

Sato, H., Yokoyama, M., and Toh, H. (2013). Genomics and computational science for virus research. *Front. Microbiol.* 4:42. doi: 10.3389/fmicb.2013.00042

Skowronski, D. M., Sabaiduc, S., Chambers, C., Eshaghi, A., Gubbay, J. B., Krajden, M., et al. (2016). Mutations acquired during cell culture isolation may affect antigenic characterisation of influenza A(H3N2) clade 3C.2a viruses. *Euro Surveill.* 21:30112. doi: 10.2807/1560-7917.ES.2016.21.3.30112

Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D., et al. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science* 305, 371–376. doi: 10.1126/science.1097211

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tsuchiya, E., Sugawara, K., Hongo, S., Matsuzaki, Y., Muraki, Y., Li, Z. N., et al. (2002). Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *J. Gen. Virol.* 83(Pt 5), 1137–1146. doi: 10.1099/0022-1317-83-5-1137

Wang, C. C., Chen, J. R., Tseng, Y. C., Hsu, C. H., Hung, Y. F., Chen, S. W., et al. (2009). Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18137–18142. doi: 10.1073/pnas.0909696106

Wang, W., Lu, B., Zhou, H., Suguitan, AL Jr, Cheng, X., Subbarao, K., et al. (2010). Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *J. Virol.* 84, 6570–6577. doi: 10.1128/JVI.00221-10

Wanzeck, K., Boyd, K. L., and McCullers, J. A. (2011). Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. *Am. J. Respir. Crit. Care Med.* 183, 767–773. doi: 10.1164/rccm.201007-1184OC

Webster, R. G., and Laver, W. G. (1980). Determination of the number of nonoverlapping antigenic areas on Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. *Virology* 104, 139–148. doi: 10.1016/0042-6822(80)90372-4

Wiley, D. C., and Skehel, J. J. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu. Rev. Biochem.* 56, 365–394. doi: 10.1146/annurev.bi.56.070187.002053

Wiley, D. C., Wilson, I. A., and Skehel, J. J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289, 373–378. doi: 10.1038/289373a0

Yokoyama, M., Naganawa, S., Yoshimura, K., Matsushita, S., and Sato, H. (2012). Structural dynamics of HIV-1 envelope Gp120 outer domain with V3 loop. *PLoS ONE* 7:e37530. doi: 10.1371/journal.pone.0037530

Yokoyama, M., Nomaguchi, M., Doi, N., Kanda, T., Adachi, A., and Sato, H. (2016). In silico analysis of HIV-1 Env-gp120 reveals structural bases for viral adaptation in growth-restrictive cells. *Front. Microbiol.* 7:110. doi: 10.3389/fmicb.2016.00110

Yuan, Y., Yokoyama, M., Maeda, Y., Terasawa, H., Harada, S., Sato, H., et al. (2013). Structure and dynamics of the gp120 V3 loop that confers noncompetitive resistance in R5 HIV-1(JR-FL) to maraviroc. *PLoS ONE* 8:e65115. doi: 10.1371/journal.pone.0065115

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Yokoyama, Fujisaki, Shirakura, Watanabe, Odagiri, Ito and Sato. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 30 November 2016
doi: 10.3389/fmicb.2016.01907

Detection of Inter-Lineage Natural Recombination in Avian Paramyxovirus Serotype 1 Using Simplified Deep Sequencing Platform

*Dilan A. Satharasinghe[1,2], Kavitha Murulitharan[1], Sheau W. Tan[1], Swee K. Yeap[1], Muhammad Munir[3], Aini Ideris[1,4] and Abdul R. Omar[1,4]**

[1] *Laboratory of Vaccine and Immunotherapeutic, Institute of Bioscience, Universiti Putra Malaysia, Serdang, Malaysia,* [2] *Faculty of Veterinary Medicine and Animal Science, University of Peradeniya, Peradeniya, Sri Lanka,* [3] *Infection and Innate Immunity Research Group, Avian Viral Diseases, The Pirbright Institute, Surrey, UK,* [4] *Faculty of Veterinary Medicine, Universiti Putra Malaysia, Serdang, Malaysia*

Newcastle disease virus (NDV) is a prototype member of avian paramyxovirus serotype 1 (APMV-1), which causes severe and contagious disease in the commercial poultry and wild birds. Despite extensive vaccination programs and other control measures, the disease remains endemic around the globe especially in Asia, Africa, and the Middle East. Being a single serotype, genotype II based vaccines remained most acceptable means of immunization. However, the evidence is emerging on failures of vaccines mainly due to evolving nature of the virus and higher genetic gaps between vaccine and field strains of APMV-1. Most of the epidemiological and genetic characterizations of APMVs are based on conventional methods, which are prone to mask the diverse population of viruses in complex samples. In this study, we report the application of a simple, robust, and less resource-demanding methodology for the whole genome sequencing of NDV, using next-generation sequencing (NGS) on the Illumina MiSeq platform. Using this platform, we sequenced full genomes of five virulent Malaysian NDV strains collected during 2004–2013. All isolates clustered within highly prevalent lineage 5 (specifically in lineage 5a); however, a significantly greater genetic divergence was observed in isolates collected from 2004 to 2011. Interestingly, genetic characterization of one isolate collected in 2013 (IBS025/13) shown natural recombination between lineage 2 and lineage 5. In the event of recombination, the isolate (IBS025/13) carried nucleocapsid protein consist of 55–1801 nucleotides (nts) and near-complete phosphoprotein (1804–3254 nts) genes of lineage 2 whereas surface glycoproteins (fusion, hemagglutinin-neuraminidase) and large polymerase of lineage 5. Additionally, the recombinant virus has a genome size of 15,186 nts which is characteristics for the old genotypes I–IV isolated from 1930 to 1960. Taken together, we report the occurrence of a natural recombination in circulating strains of NDV in commercial poultry using NGS methodology. These findings will not only highlight the potential of RNA viruses to evolve but also to consider the application of NGS in revealing the genetic diversity of these viruses in clinical materials. Factors that drive these evolutionary events and subsequent impact of these divergences on clinical outcome of the disease warrant future investigations.

Keywords: Newcastle disease virus, avian paramyxovirus 1, next-generation sequencing, phylogenetic analysis, recombination

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Oscar Negrete,
Sandia National Labs, USA
Takemasa Sakaguchi,
Hiroshima University, Japan
Kaoru Takeuchi,
University of Tsukuba, Japan

****Correspondence:***
Abdul R. Omar
aro@upm.edu.my

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *19 September 2016*
Accepted: *15 November 2016*
Published: *30 November 2016*

Citation:
Satharasinghe DA, Murulitharan K, Tan SW, Yeap SK, Munir M, Ideris A and Omar AR (2016) Detection of Inter-Lineage Natural Recombination in Avian Paramyxovirus Serotype 1 Using Simplified Deep Sequencing Platform. Front. Microbiol. 7:1907. doi: 10.3389/fmicb.2016.01907

INTRODUCTION

Newcastle disease virus (NDV) is a type species of an avian paramyxovirus serotype 1 (APMV-(1), which belongs to genus *Avulavirus* in the family *Paramyxoviridae*. The virus carry negative-sense, single-stranded, and non-segmented RNA genome, which encodes for at least six structural proteins, including nucleocapsid protein (NP), phosphoprotein (P), matrix protein (M), fusion protein (F), haemagglutinin-neuraminidase (HN), and large protein (L) (Chambers et al., 1986). Through the RNA editing of the P gene, one accessory non-structural protein (known as V) is synthesized, and probability a second protein named W is also produced (Locke et al., 2000). The HN and F proteins are surface glycoproteins that determine the virus neutralization and protection. The NP, P, and L proteins encapsidate the viral nucleic acid; the M protein lined the virus envelope and V protein interfere with interferon (IFN) responses (Dortmans et al., 2011; Zenglei and Liu, 2015).

Depending on the pathogenicity in chicken, NDV strains are classified into highly virulent (velogenic), intermediate (mesogenic), and avirulent (lentogenic) strains (Alexander, 1997). The amino acid sequence of the F protein cleavage site has been used to determine the pathogenicity of NDV (Nagai et al., 1976). Generally, the sequence of the F protein cleavage site in mesogenic and velogenic strains of NDV consists of a polybasic cleavage site (R/K) RQ(R/K) R↓F, which is readily recognized by ubiquitous furin, an intracellular protease abundant in several cells and tissues, and consequently cleaves in many tissues resulting in systematic infection. In contrast, initiation of viral infection by cleavage of the F protein in lentogenic strains is restricted to respiratory and enteric systems and thus limiting the infections in these organs (Collins et al., 1993).

Although all NDV's strains are grouped within APMV-1, yet viruses show significant genetic diversity (Alexander, 1997; Aldous et al., 2003; Kim et al., 2007a). Without any incongruity, there are two methods which are currently being used to classify APMV-1 strains into different clusters. Aldous et al. (2003) have divided APMV-1 into six lineages and 13 sub-lineages, which later included three sub-lineages (Aldous et al., 2003; Snoeck et al., 2009). According to the second system, APMV-1 has been divided into two main categories which are represented as class I and class II. The class I is also subdivided into nine genotypes while the class II is subdivided into at least 13 genotypes (Ballagi-Pordány et al., 1996; Czeglédi et al., 2006; Kim et al., 2007b). In general, there is just slight dissimilarity that exists amid the two systems (Miller et al., 2010) in which lineage 6 comprised class I viruses and lineage 1, 2, 4, and 5 consisted of viruses belonging to genotype I, II, VI, and VII of class II, respectively. Furthermore, highly diverse lineage 3 comprised of genotype III, IV, V, and VIII of class II viruses (Miller et al., 2010). Several previous studies have shown that genotype V, VI, and VII of class II viruses were the most prevalent viruses presently circulating globally. Out of these, lineage 5 or genotype VII of class II viruses were the most predominantly isolated strains of NDV from Asia, Africa, Middle East, and South America (Khan et al., 2010; Miller et al., 2010; Munir et al., 2012; Perozo et al., 2012). A recent study has classified NDV strains into several new genotypes based on the mean inter-populationary evolutionary distance of 10% as the cutoff value (Diel et al., 2012). Based on that study, viruses in class I formed a single genotype, while class II viruses were separated into 15 genotypes including 10 earlier established genotypes (I - IX and XI) and five new genotypes (X, XII, XIII, XIV, and XV). The subsequent study introduced two additional genotypes into this classification system (XVII and XVIII) from NDV outbreaks occurring in the Asia, west, and central Africa (Munir et al., 2012; Snoeck et al., 2013). Taken together, none of the classification systems represent a clear lineage/genotype-specific disease potential, geographical distribution or host specifications.

Similar to other RNA's viruses, NDV holds potential for evolution (Miller et al., 2010). However, limited information is available that clearly defines events of evolution, recombination, and possible gained pathogenicities of NDV strains in avian hosts. This lack of information is partly due to (i) mainly NDVs are characterized based on the partial sequence of F or HN genes which are insufficient to identify recombination (Song et al., 2011), (ii) full genetic characterized of majority of APMVs is performed using conventional methodologies (PCRs and Sanger sequencing) which is not only error prone but also less capable to identify poorly presented virus population in the complex samples, and (iii) initial clinical material appeared to be insufficient for amplification of full genome in conventional PCRs. This requires propagation of viruses in chicken eggs, and this process may favor the replication of fitter viruses and masking the recombined strains that by any means are lower in population. Despite these shortcomings, recombination events, and evolution diversities have been reported in the NDV (Qin et al., 2008; Chong et al., 2010). Previously, Han et al. (2008) have reported a natural recombination NDV isolate named cockatoo/Indonesia/14698/90 (AY562985) with anonymous major parental lineage and two minor parental-like lineages derived from vaccine lineage and anhinga/U.S. (Fl)/44083/93 lineage, in that order (Han et al., 2008).

One of the driving forces for recombination and increasing pathogenic potential of viruses is the immune pressure imposed by the extensive vaccination (Read et al., 2015). For NDV, the primary means of disease control has been the employment of exhaustive vaccination program in both layer and broiler poultry sectors. Despite mass vaccination being practiced, NDV outbreaks have been reported throughout the world and have led to substantial economic losses to the industry over the years (Miller et al., 2010). This escaped protection has been linked to the partial protection by the current vaccines and persistent virus shedding in the immunized birds (Xiao et al., 2012; Samuel et al., 2013).

In Malaysia, an intensive vaccination program has been placed, however, the disease is continuously emerging, and the impact is enormous (Yusoff and Tan, 2001; WAHID, 2015). In this regards, we have recently characterized several isolates of NDV from various states of Malaysia from 2004 to 2005 (Tan et al., 2010). This characterization was based on the partial F gene sequences which limit the full genetic characterization and detection of recombination events. Here, we optimized a simple, robust and less resource-demanding methodology for the whole genome sequencing of APMVs using next-generation

sequencing (NGS) technology. Using this protocol, we generated unbiased consensus-level full genomes of five APMVs strains which were obtained from clinical outbreaks during 2004–2013 in Malaysia. Phylodynamics and evolutionary analysis revealed recombination between lineage 2 and 5 and shown features that are unique for this lineage/genotype of APMV-1. The presented results are crucial to establishing bases for the vaccine-induced immune pressures, vaccine failure, and potential of APMVs to evolve for higher pathogenicity.

MATERIALS AND METHODS

Samples Collection

Five NDV isolates were obtained from individual outbreaks in the commercial poultry farms that were vaccinated for Newcastle disease (ND) located in different states in Malaysia (**Table 1**). Two isolates, IBS002/11 and IBS005/11, were isolated in 2011 while other three isolates, MB128/04, MB 076/05, and IBS025/13, were isolated in 2004, 2005, and 2013, respectively. Apart from IBS025/13, all samples were detected positive previously using reverse transcriptase PCR (RT-PCR) which is based on the partial F gene (data not shown). In addition, these isolates have been characterized in the past as genotype VII NDV isolates which are based on the partial F gene sequencing (Berhanu et al., 2010; Tan et al., 2010; Roohani et al., 2015).

Virus Specimens

All NDVs used in this study were triple plaque-purified on chicken embryo fibroblast (DF1 cells) and propagated by inoculating into 9 days-old specific-pathogen-free (SPF) embryonated chicken eggs and stored in liquid nitrogen. Virus stocks of the selected NDV isolates were thawed and propagated in the allantoic cavity of 9 days old SPF embryonated chicken eggs according to European Community Directive 92/66/EC (CEC, 1992) and identified using hemagglutination (HA) test. Allantoic fluids from the sample showing high HA titers more than 2^7 were divided into working stocks and stored at −20°C. These allantoic fluids were used as working stocks for pathogenicity assessments and sequencing.

RNA Extraction and Quality Control

Genomic viral RNA was extracted from the infected allantoic fluid of SPF eggs by using TRIzol® Reagent (Invitrogen, USA) according to the manufacturer's instructions. After air drying, the RNA pellet was resuspended in 40 μL nuclease-free water (Ambion, USA). The quantity and purity of the extracted RNA were checked by Eppendorf BioSpectrometer® subsequently, and was then stored at −80°C for further analysis.

Designing of Primers

Based on the complete genome of NDV strain chicken/Banjarmasin/010/10 (HQ697254.1), five pairs of primer sets (**Table 2**) were designed to amplify the whole genome of five Malaysian NDV isolates. The full genome of NDV was divided into 5 fragments encompassing hypervariable regions flanked by more conserved sequence regions. The fragments that were identified with the most conserved regions were subjected to NCBI Primer 3 and Primer-BLAST tool and five diverse primer pairs were designed. An additional primer pair was designed for the third fragment, which consists of the highly variable regions.

cDNA Synthesis

First-strand cDNA synthesis (reverse transcription) was performed using MMLV Reverse Transcriptase First-Strand cDNA synthesis kit (Epicentre, USA). Briefly, the following components were combined on ice: 6.5 μL of RNase-Free Water, 5 μL of viral RNA sample and 1.5 μL of fragment 1 forward primer (**Table 2**) for a 12.5 μL total reaction volume, incubated at 65°C for 2 min with heated lid and chilled on ice for 1 min. A second reagent mix of 2 μL MMLV RT 10 × Reaction Buffer, 2 μL 100 mM DTT, 2 μL dNTP PreMix, 0.5 μL RiboGuard RNase Inhibitor and 1 μL MMLV Reverse Transcriptase were added to each first-strand cDNA synthesis reaction and mixed gently on the ice. The reaction was incubated at 37°C for 60 min followed by heating at 85°C for 5 min and chilled on ice for at least 1 min. The cDNA was briefly centrifuged and immediately used or stored at −80°C for future analysis.

TABLE 1 | Clinical description and vaccination history of Malaysian NDV isolates used in this study.

NDV isolate	Year	Origin	Farm type	Case history	Vaccination history	References
MB128/04	2004	Selangor	Broiler	Not available	Not available	Berhanu et al., 2010
MB076/05	2005	Sabah	Broiler	Depression, marked conjunctivitis, anorexia, and facial edema	Not available	Berhanu et al., 2010
IBS002/11	2011	Johor	Broiler	The infection started at day 14 with 5% mortality. Bird showed signs of coughing, sneezing, and torticollis	At day 8 and 14 with live ND+IB	Roohani et al., 2015
IBS005/11	2011	Penang	Broiler	The infection started at day 29 with 5% mortality. Depression and greenish diarrhea, respiratory rales	At day 2, 8, and 12 with ND+IB	Roohani et al., 2015
IBS025/13	2013	Penang	Broiler	The infection started at day 18 with 4% mortality. Bird showed signs of coughing, sneezing, greenish diarrhea and torticollis	At day 14 with ND+IB	This study

ND, Newcastle disease; IB, Infectious bronchitis.

TABLE 2 | Primers used to amplify the complete genome of lineage 5 NDV.

Fragment	Location	Forward primer sequence (5′ → 3′)	Reverse primer sequence (5′ → 3′)	Size
1	1–2792	ACCAAACAGAGAATCTGTGAGGTAC	TCGGGCTACTGCTCGGAGAT	2792
2	2541–5926	GTCCAGCTACCTGCCGACTT	GCCTGTCACGATGACTTGAGA	3386
3*	5673–9156	CTGAAGGCGCACTCACTACG	AGATGTGCCGCCGTAGAAGAT	3484
	4258–8657	CCTGCGGAGTGTGAAAGTCA	GGAGCACTCCGGTTATCTTGG	4400
4	8644–12,183	AACCGGAGTGCTCCATCCC	AGATTCCCAGCTGTGGGCAG	3540
5	11,745–15,192	CCCGGAACAGAAGCTGGTCA	ACCAAACAGAGATTTGGTGAACGAC	3448

**An additional primer was designed for fragment 3 which was located in the hypervariable region of the genome. Primer locations are based on NDV strain chicken/Banjarmasin/010/10 (HQ697254.1).*

Synthesis of Double Strand DNA

Based on the manufacturer's instructions of using KAPA HiFi Hot Start Ready Mix (KAPA Biosystems, USA) kit, the cDNA was used to synthesize a double strand DNA. Briefly, 3 μL of cDNA was combined with 12.5 μL of 2 × KAPA HiFi Hot Start Ready Mix, 0.9 μL of 10 μM forward primer and reverse primer up to 25 μL volume with PCR-grade water. The thermal cycling protocol for the PCR was included with an initial denaturation at 95°C for 3 min followed by a sequence of 30 cycles, which comprised denaturation at 98°C for 20 s, annealing at 60°C for 20 s, extension at 72°C for 2 min and final extension hold at 72°C for 5 min. The PCR product was detected in 0.8% agarose gel electrophoresis and purified using MEGAquickspin™ Total Fragment DNA Purification Kit (iNtRON Biotechnology, South Korea) as per manufacturer's instructions.

Next-Generation Sequencing Library Preparation and Sequencing

Each DNA originated from the isolates was quantified by using Qubit dsDNA HS Assay Kit (Invitrogen, USA) and then normalized to 0.2 ng/μL. A total of 1 ng of DNA was subjected to library preparation using Nextera XT DNA Sample Prep Kit (Illumina Inc., San Diego, CA, USA) by following the manufacturer protocol. Briefly, the DNA was tagmented (fragmented and tagged) by the Nextera XT transposome. The tagmented DNA was used as template in a 50 μl of PCR with 12 cycles and processed as outlined in the Nextera XT protocol. Additionally, AMPureXP beads (Beckman Coulter Inc., Fullerton, CA, USA) was applied to purify the amplified DNA.

After PCR clean up, DNA fragment size and library concentration was analyzed by using 2100 Bioanalyzer (Agilent Technologies, USA) and Library Quantification Kit (KAPA Biosystems, USA). Afterward, DNA libraries were normalized to 4 nM and libraries with unique indexes were pooled in equal volumes. Pooled libraries were denatured and diluted with 0.2 N NaOH and pre-chilled Hybridization Buffer (HT1) to produce a denatured 12 pM library in 1 mM NaOH solution. The final library was sequenced using MiSeq (Illumina Inc., San Diego, CA, USA) with the read length of 2 × 150 bp.

Analysis of the Data

The overlapping paired-end reads were filtered on the Phred quality score (Q30) and imported to CLC Genomics Workbench software version 7.5.1 (CLC bio, Aarhus, Denmark) for adapter trimming and *de novo* assembly of the paired-end reads to contigs. The contigs were then subjected to BLASTN at NCBI and based on the highest sequence similarity and lowest *E*-value, reference genomes were determined. Low coverage contigs were excluded and when necessary, partial but overlapping contigs were combined. Final consensus having more than 99% coverage to full genome sequences were then examined for appropriate assembly depending on the length and the presence of the expected intact NDV open reading frames.

Confirmation Using Sanger Sequencing

Two gaps were observed in the consensus of IBS005/11 at the nucleotide positions of 1708–1743 and 1821–1882. Forward primer sequence 5′-GCCATCCCAAGACAACGACA-3′ aligning nucleotide position of consensus at 1555 and the reverse primer sequence 5′- CCCTGGGCCGTTATTATGCT-3′ aligning position of IBS005/11 consensus at 1956 were designed by using NCBI Primer 3 software to close the gap. The PCR product was observed in 1.5% agarose gel. Similarly, 7 nucleotides gap was observed in the consensus of MB128/04 at the nucleotide positions 1698–1708. This gap was closed using the same forward and the reverse primer sequences earlier mentioned. Additionally, the recombination region (1–2766 bp) including the 6 nucleotides deletion at the position 1647 of 5′ NCR of the NP gene was validated twice by sequencing the fragment 1 (**Table 2**) of the IBS025/13 NDV isolate by primer walking technology.

The cDNA template was synthesized as previously described and PCR was performed according to PCR conditions mentioned above. Both PCR products were purified with MEGAquickspin™ Total Fragment DNA Purification Kit (iNtRON Biotechnology, South Korea) as per manufacturer's instructions. Sanger sequencing was performed from both directions in order to close the gaps and confirmation of recombination region in the IBS 025/13 isolate. Finally, the obtained sequences were aligned with the consensus by using MEGA6 software (Tamura et al., 2013).

Amplification of the 3′- and 5′- Terminal Ends of the Viral RNA

To determine the leader and trailer at the 3′- and 5′- terminal ends of the viral RNA, rapid amplification of

cDNA ends (RACE) was performed as described earlier on de Leeuw and Peeters (1999). The 5′ end primer L-cDNA (5′- AAGTCACAATACTGGGTCTCAG -3′), which was designed from L gene of genotype VII NDV strain, chicken/Banjarmasin/010/10 (HQ697254.1) was used to generate single-stranded cDNA as described above. The T4 RNA ligase (New England Biolabs Inc, USA) was then used to ligate the generated single-stranded cDNA so as to anchor primer (5′- CACGAATTCACTATCGATTCTGGATCCTTC -3′). One micro liter of ligation mixture was used for PCR with KAPA HiFiHotStartReadyMix (KAPA Biosystems, USA) in accordance with the manufacturer's instructions. The primers used were anchor-complementary (5′-GGATCCAGAAT-CGATAGTGAATTCG-3′) and L-PCR (5′-CAGCCAAGGGAT-ATTACAGTAACT-3′). The PCR consisted of 30 cycles of 1 min at 98°C, 20 s at 60°C and 20 s at 72°C. The PCR products were then cloned into pJET 1.2 blunt/cloning vectors (Thermo Fisher Scientific, Inc, USA) and verified by sequencing. The anchor primer was ligated to the 3′ prime end of genomic RNA with T4 RNA ligase as described above in order to determine the sequence of the 3′-terminal ends. Then, the anchor-complementary primer was used to generate the cDNA. Hence, anchor-complementary and NP-PCR primers were used to conduct PCR. Primer (5′-GGAGCTGCTCGTATTCGTC-3′) was used to generate fragments for strains MB076/05, IBS002/11, IBS005/11, and MB128/04. Meanwhile, NP-PCR primer (5′-CGAGGAGCTGTTCGTACTCATCAA-3′) was used for strain IBS025/13. The PCR conditions were the same as have been described above.

Recombination among NDV Sequences

SimPlot program (Ray, 2003) was used to identify putative recombination breakpoints in the sequenced whole genome of NDV isolates and to identify sequences possibly originated from a recombination. The program is based on a sliding window process and consists of a way of graphically displaying the coherence of the sequence relationship over the entire length of a set of aligned homologous sequences. The window width and the step size were set to 200 bp (or 500 bp) and 20 bp, respectively. In addition, Recombinant Program v4.56 (RDP4) was used for the detection of recombination events, likely parental isolates of recombinants and recombinant breakpoints. Various methods such as RDP, GENECONV, Chimaera, MaxChi, BOOTSCAN, and SISCAN with default settings (Martin et al., 2005) were used by the RDP4 program.

Phylogenetic and Pairwise Sequence Comparisons (PASC) Analysis

The nomenclature of the NDV isolates in this study was based on genotypes (Ballagi-Pordány et al., 1996; Czeglédi et al., 2006) and lineages (Aldous et al., 2003; Snoeck et al., 2009; Munir et al., 2012). CLUSTALW (Thompson et al., 1997) was used to align all the NDV sequences while MEGA6 software was used to analyze all the phylogenetic. The model with lowest BIC (Bayesian Information Criterion) value was selected as the most suitable model for phylogenetic analysis (Tamura et al., 2013). The reliability of the lineages defined for NDV was determined by Pair-Wise Sequence Comparisons (PASC) using selected NDV strains available in GenBank representing all lineages except lineage 6 and Malaysian isolates. All reference isolates used in this study were named first with accession number followed by country of origin, respected lineage and year of isolation. Also, mean distances among and within lineages were calculated using PASC in MEGA6 software.

Ethics Statement

The animals used for this research were kept in the house and animal care procedures were conducted according to the local animal welfare regulations and EU directive (EU Directive on the protection of animals used for scientific purposes 2010/63/EU) (EU, 2010) under bio-safety level (BSL2) enhanced experimental animal facility at the Faculty of Veterinary Medicine, UPM. The animal experimental protocol for Intracerebral Pathogenicity Index (ICPI) study was approved by the Institutional of Animal Care and Use Committee (IACUC), Faculty of Veterinary Medicine, UPM (reference number UPM/IACUC/AUP-R028/2013); the local animal care authority. Animals were monitored for a minimum of three times per day by a qualified and registered veterinarian to ensure animal welfare and health. The end point of all these *in vivo* animal experiments was death; although, a humane endpoint was pre-defined in the protocol and applied to prevent any pain, distress or suffering. The humane endpoint was decided for chicks used for ICPI study manifested terminal clinical signs. Animals showing terminal signs including anorexia and paralysis were sacrificed by cervical dislocation under sedation in accordance with standard guidelines (EU, 2010; Leary et al., 2013).

Intracerebral Pathogenicity Index Assessment in SPF Chickens

Pathogenicity of NDV isolates was assessed by using a standard ICPI test as earlier stated (OIE, 2012). Briefly, 50 μl of allantoic fluid of each NDV isolate with an HA titer more than 16 HA units/50 μl and diluted 10-fold in PBS without antibiotics were inoculated intracerebrally to 1-day-old chicks ($n = 10$). Ten chicks were kept as uninoculated control. All the chicks were being monitored at least three times per day for an 8-day period of observation. The NDV isolates that scored an ICPI > 1.5 were identified as velogenic strains, <0.7 as lentogenic and those with intermediate ICPI values were considered as mesogenic strains (OIE, 2012).

RESULTS

Whole Genome Sequencing of NDV

The present study proposes a pipeline to generate consensus level sequences with more than 99% coverage of the full genome. The consensus level genome sequences of five NDV isolates belonging to genotype VII were generated for the first time in Malaysia. The whole genome sequences of these NDV isolates were deposited in the GenBank and are available under the following accession numbers: KR074407 for MB128/04, KR074406 for

TABLE 3 | Summary of the genome organization and the predicted protein length of the sequenced MB128/04, MB076/05, IBS002/11, and IBS005/11 isolates.

Gene	Gene start position (nt)	3′UTR length (nt)	Coding sequence positions (nt)	5′UTR length (nt)	Gene end position (nt)	Intergenic region length (nt)	Gene length (nt)	Protein length (aa)
NP	56–65	66	122–1591	217	1799–1808	1	1753	489
P	1810–1819	83	1893–3080	180	3251–3260	1	1451	395
M	3262–3271	34	3296–4390	112	4493–4502	1	1241	364
F	4504–4513	46	4550–6211	84	6284–6295	31	1792	553
HN	6327–6336	91	6418–8133	195	8319–8328	47	2002	571
L	8376–8385	11	8387–15,001	77	15,069–15,078	–	6703	2204

UTR, untranslated region.

TABLE 4 | Summary of the genome organization and predicted protein lengths in the sequenced IBS025/13 isolate.

Gene	Gene start position (nt)	3′UTR length (nt)	Coding sequence positions	5′UTR length (nt)	Gene end position (nt)	Intergenic region length (nt)	Gene length (nt)	Protein length (aa)
NP	56–65	66	122–1591	210	1791–1801	2	1746	489
P	1804–1813	83	1887–3074	180	3245–3254	1	1451	395
M	3256–3265	34	3290–4384	112	4487–4496	1	1241	364
F	4498–4507	46	4544–6205	84	6280–6289	31	1792	553
HN	6321–6330	91	6412–8127	195	8313–8322	47	2002	571
L	8370–8379	11	8381–14,995	77	15,063–15,072	–	6703	2204

UTR, untranslated region.

MB076/05, KR074404 for IBS002/11, KR074405 for IBS005/11 and IBS025/13 for KT355595.

Summary of the genomic features of the Malaysian isolates, including starting and ending positions of each gene, intergenic, and coding regions, represented in **Tables 3, 4**. The genome length of 4 NDV isolates (MB128/04, MB076/05, IBS002/11, and IBS005/11) and IBS025/13 isolate was calculated to be 15,192 nts and 15,186 nts respectively, which putatively consisted of 6 different genes with each gene containing a single or multiple open reading frames (ORFs) as indicated in **Tables 3, 4**.

TABLE 5 | Biological and molecular characterization of NDV isolates.

NDV isolate	F cleavage site	ICPI
MB128/04	^{112}RRQKR↓F^{117}	1.68
MB076/05	^{112}RRQKR↓F^{117}	1.84
IBS002/11	^{112}RRRKR↓F^{117}	1.76
IBS005/11	^{112}RRRKR↓F^{117}	1.71
IBS025/13	^{112}RRQKR↓F^{117}	1.63

Biological and Molecular Pathogenicity Assessment

The isolates in the current study exhibited ICPI values more than 1.5 and the F protein cleavage site sequences exhibited ^{112}R/K-R-Q-R/K-R↓F^{117} motif similar to mesogenic and velogenic NDV strains, which is another characteristic considered by the World Organization for Animal Health (OIE) to assess the pathogenicity (OIE, 2012) (**Table 5**).

Phylogenetic and Pairwise Sequence Comparison (PASC) Analysis of NDV Isolates

Genomic information is the basis used by the two currently available methods in NDV classification, where one is to identify strains by lineages and the other by classes/genotypes. The phylogenetic and pairwise sequence comparison data presented in this study is based on the lineage classification proposed by Aldous et al. (2003) and Munir et al. (2012). According to the phylogenetic analysis of partial F gene, all the isolates clustered with lineage 5 (**Figure 1**). Moreover, in line with the PASC analysis, the isolates also have the highest percentage identity (PI) to NDV strains belonging to lineage 5. The NDV isolates MB128/04 and MB076/05 isolated in 2004 and 2005, respectively, showed the highest PI of 92.41% to lineage 5 (**Table 6**). Meanwhile, PASC analysis of the genetic similarity among sub-lineage under lineage 5 indicated that the isolates showed the highest PI to sub-lineage 5a. However, except for IBS025/13, the genomic divergence between the lineage 5a and the sequenced viruses isolated during the period from 2004 to 2011 is increasing (**Table 7**). The phylogenetic tree was drawn based on nucleotide sequence of complete genome sequences, and it shows the clustering of the isolates with lineage 5 (**Figure 2A**). The robustness of the genetic groupings and topology of the phylogenetic trees were also confirmed by the findings obtained from the PASC analysis of the complete genome as well as F, HN,

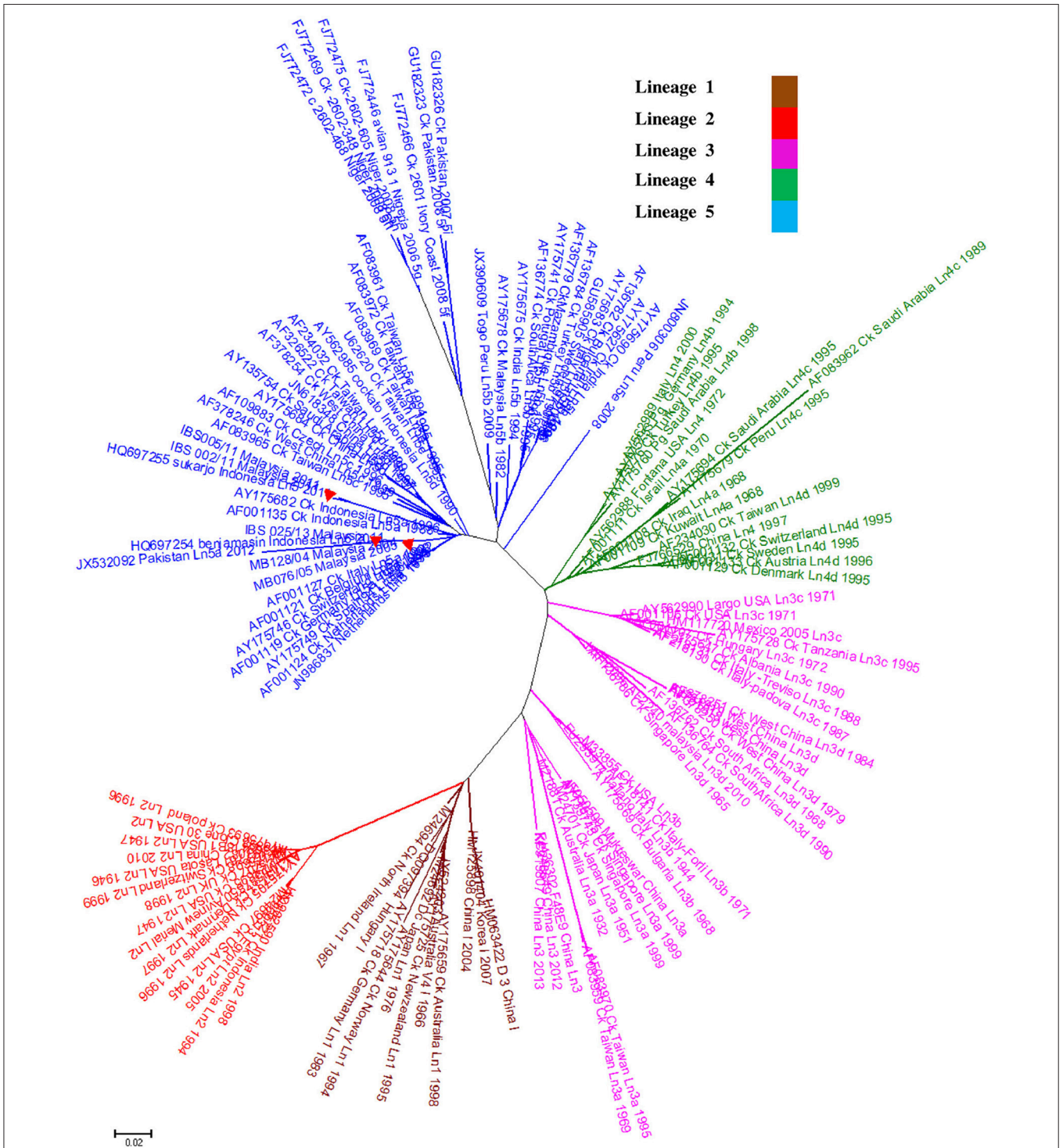


FIGURE 1 | A phylogenetic analysis of the Newcastle disease viruses; Phylogenetic relationship among 120 published NDV isolates and 5 Malaysian isolates based on F gene nucleotide sequences between 47 and 422 positions. The sequences used were obtained from GenBank. As identified by Aldous et al. (2003), Class II viruses representing lineages 1–5 ($n = 120$) and Malaysian isolates from 2004 to 2013 ($n = 5$) identified with 373 base pair region encoding the amino terminal end of the fusion protein. Tree construction was done using the Neighbor-Joining method with the maximum composite likelihood substitution model after 1000 bootstrap replication by using MEGA 6. Lineage 6 of APMV-1 was excluded in the analysis. Isolates presented in this study was marked with a red triangle (▲).

TABLE 6 | Pairwise sequence comparison of the partial F gene sequences of Malaysian isolates based on different lineages.

NDV Isolate	Percentage identity (%)				
	1	2	3	4	5
MB128/04	83.74	81.16	87.42	89.09	92.41
MB076/05	83.74	81.16	87.42	89.09	92.41
IBS002/11	81.48	79.84	84.65	85.81	88.42
IBS005/11	81.48	79.84	84.65	85.81	88.42
IBS025/13	83.57	80.66	86.50	87.87	91.32

TABLE 7 | PASC analysis of the partial F gene sequences of Malaysian isolates based on lineage 5.

NDV isolate	Percentage identity (%)								
	5a	5b	5c	5d	5e	5f	5g	5h	5i
MB128/04	96.22	92.23	93.74	93.51	93.16	85.79	86.33	83.56	86.33
MB076/05	96.22	92.23	93.74	93.51	93.16	85.79	86.33	83.56	86.33
IBS002/11	91.51	87.57	90.88	90.03	88.67	81.77	82.31	79.54	83.51
IBS005/11	91.51	87.57	90.88	90.03	88.67	81.77	82.31	79.54	83.51
IBS025/13	94.82	91.37	92.58	92.6	91.89	84.45	84.46	81.68	85.66

M, and L protein sequences of MB128/04, MB076/05, IBS002/11 and IBS005/11 isolates, respectively, whilst PASC analysis of NP and P protein sequences of IBS025/13 isolate showed the highest percentage identity to lineage 2 viruses (**Table 8**). Furthermore, maximum likelihood phylogenetic trees constructed based on NP and P proteins of IBS025/13 isolate indicated that those proteins were clustered with lineage 2 viruses and substantiated the results obtained from PASC analysis (**Figure 2B**).

Recombination Analysis of NDVs Using Simplot and RDP

A standard similarity plot (Simplot) (Ray, 2003) was used to analyze the possible events of recombination in the sequenced isolates with representative isolates of lineages 1–5 obtained from GenBank. Since the analysis revealed that NP and P gene regions of isolate IBS025/13 have the highest genomic similarity toward lineage 2 viruses over lineage 5 (**Figures 3A,B**), we also investigated the first 1/3 of NDV genome alignment to identify potential breakpoints by using recombinant detection program (RDP) (Martin et al., 2005). Possible breakpoints at nucleotide position 1–2766 and 2767–2984 were detected by *p*-value equal to RDP (8.418E-4), Geneconv (1.617E-3), BootScan (8.349E-4), MaxChi (1.612E-1) and SiScan (2.509E-10). The most convincing evolutionary evidence for recombination was the occurrence of incongruent phylogenetic trees (Zhang et al., 2010). Maximum composite likelihood phylogenic trees were constructed in both break points using MEGA 6 software with 1000 bootstrap values according to the best model selected on lowest BIC. The results confirmed clustering of IBS025/13 isolate with lineage 2 NDVs at breakpoint 1–2766 and breakpoint 2767–2984 with lineage 5 NDVs (**Figures 4A,B**). Further evidence on the event of recombination in NDV was provided by sequence comparisons within the species-defining clusters (Han et al., 2008). Nucleotide fragment of 1–2766 of IBS025/13 was used as a query in BLAST in the GenBank, which showed the highest sequence identity of 99% to Lasota (AF077761), B1 (AF309418), Clone 30 (Y18898) lineage 2 NDV vaccine strains (data not shown). Sequence analysis confirmed the NP gene of IBS025/13 from nucleotide position 55–1801, which include both the 3′ and 5′ untranslated regions (UTR) were derived from lineage 2 isolates (**Table 4**). In addition, the deletion of 6 nts at the UTR region of NP gene revealed similar genome length to lineage 2 NP genes. Meanwhile, the entire 3′ UTR of P gene (1804–1886) and two-thirds of P gene from position 1887–2776 of IBS025/13 also resemble lineage 2 isolates (**Table 4**).

DISCUSSION

In 2013, more than 3000 ND outbreaks were reported from Asian countries to the World Organization for Animal Health (WAHID, 2015). Characterizing and understanding the molecular epidemiology of the currently circulating NDV strains in the world, is essential for the controlling and preventing ND outbreaks (Miller et al., 2010). The current approach used only five sets of primers to cover the whole genome of NDV, and the protocol could support the multiplexing of 94 samples by tagging unique combination of indexes during a single sequencing run with a fast turn-around time, which dramatically reduced the cost and laborious work needed to be conducted in complete genome sequencing. Since adopted method used a unique combination of indexes to each NDV sample during multiplexing and raw sequenced data was filtered according to the tagged indexes, generated sequences are originated purely from the NDV isolates obtained from outbreaks. Moreover, NGS based on Ion Torrent Personal Genome Machine (PGM, Life technologies), has been used to sequence the complete genome of avian paramyxoviruses type 4 (Wang et al., 2013).

Genome lengths of the MB128/04, MB076/05, IBS002/11 and IBD005/11 isolates were similar to the genome length of genotype V–VIII. and IX in class II APMV-1. Studies have shown that the genome length of these genotypes have evolved by an insertion of 6 nts at the nucleotide position of 1647 in the 5′ non-coding region (NCR) of the NP gene of early genotypes (I–IV), which have genome length of 15,186 nts (Czeglédi et al., 2006). Interestingly, genome length of IBS025/13 NDV isolated in 2013 was 15,186 nts, similar to the genome length of early NDV genotypes (I–IV) isolated during 1930–1960s (Czeglédi et al., 2006). In this isolate, deletion of 6 nts at the same position of 5′ NCR of the NP gene was observed subsequently confirmed by targeted region amplification and sequencing using primer walking technology. In addition, the NP and P gene regions of IBS025/13 were sequenced thrice using Sanger sequencing to confirm the recombinant events.

The OIE recognizes the ICPI as a test which can be used to assess the pathogenicity of NDV. An NDV strain with an ICPI $\geq$ 0.7 is identified as a virulent or "notifiable" to the

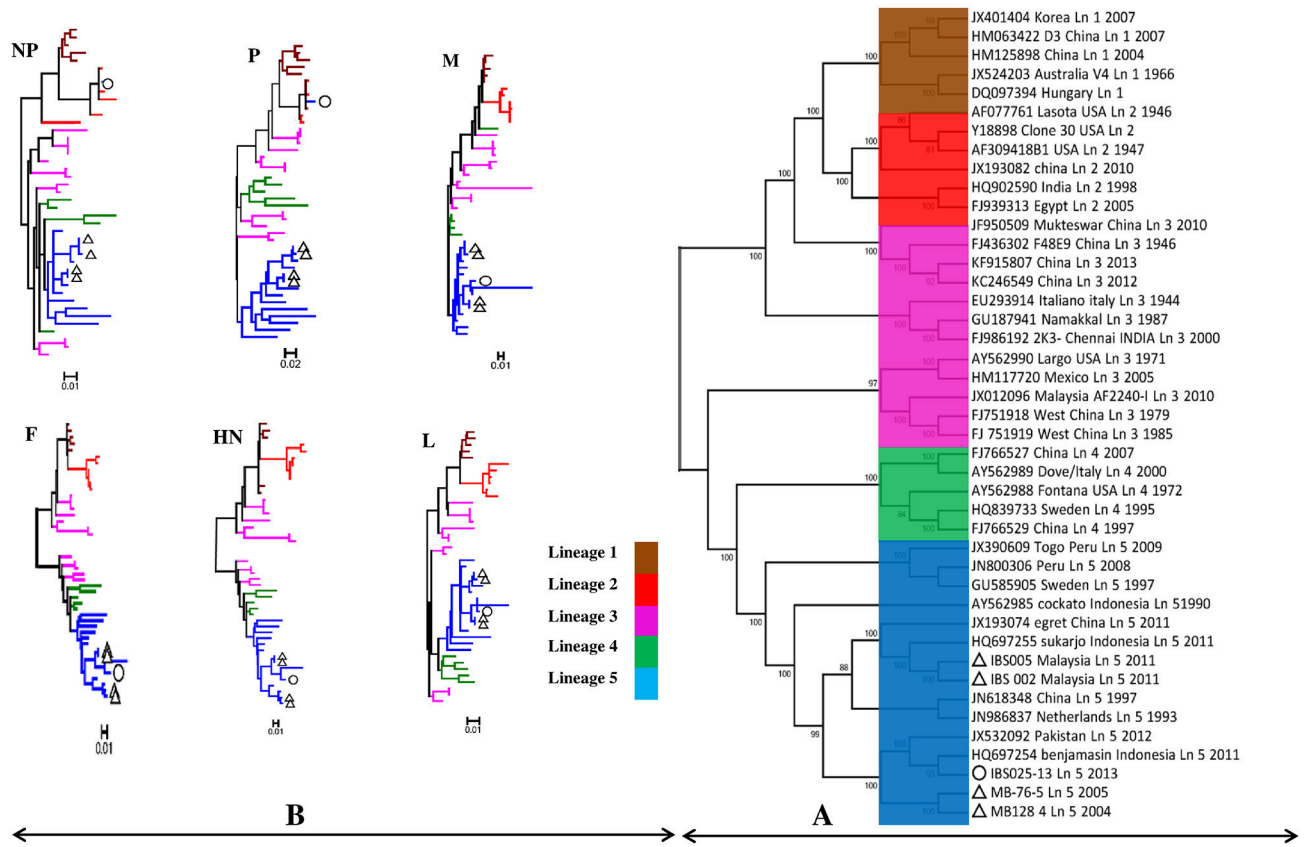


FIGURE 2 | The maximum likelihood phylogenetic trees constructed by using 5 Malaysian NDV isolates. (A) The phylogenetic tree based on nucleotide sequence of complete genome compared with 38 representative complete genomes of lineage 1–5 belongs to APMV-1. **(B)** The phylogenetic trees based on amino acids sequence of NP, P, M, F, HN, and L proteins compared with 38 representative protein sequence of lineage 1–5 belongs to APMV-1. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA6. Malaysian isolates MB128/04, MB076/05, IBS002/11, and IBS005/11 presented in this study are marked with a triangle (Δ) and IBS025/13 with a circle (○).

OIE. The cleavage site motif ^{112}R/K-R-Q-R/K-R↓F^{117} is cleavable by a wide range of proteases that subsequently cause systemic infection (Panda et al., 2004; Maminiaina et al., 2010; Miller et al., 2010). The cleavage site motif of the isolates also demonstrated the sequences of velogenic strains and these findings confirm the velogenic nature of the Malaysian NDV isolates (**Table 5**) reported in this study.

Previously, Munir et al. (2012) have indicated that NDV sub-lineages under lineage 5 can be identified with cut-off PI at 95% based on partial F gene sequences. In that study, viruses in lineage 5 can be further divided into 9 sub-lineages, from 5a to 5i with previously characterized NDV isolates including recent isolates reported from African countries and Pakistan. Based on the above study, MB128/04 and MB076/05 can be placed in the sub-lineage 5a. One of the salient findings of this study is, although IBS002/11 and IBS005/11 that were isolated in 2011 have the highest PI-value to the sub-lineage 5a, the PI-value was <95%. The findings illustrate the evolutionary pattern observed in NDV isolates from 2004 to 2011 and urged a reviewing of the current sub-lineage cut off point of lineage 5 in NDV classification or placing NDV isolates IBS002/11 and IBS005/11 under a new sub-lineage of lineage 5. Furthermore, the findings of this study support a reviewing of the current nomenclature of NDV as proposed by Miller et al. (2010) and Munir et al. (2012). The phylogenetic tree drawn based on complete nucleotide sequences of isolates shows the clustering of the isolates with lineage 5 and these findings agreed with Miller et al. (2010), who propounded that currently, the predominant circulating NDV isolates in the Asian region are of lineage 5.

In comparison to the traditional NDV classification method by partial F gene, the whole genome analysis is advantageous for the identification of recombination events. Other previous studies have reported the evidence of recombination in F gene sequences between genotype II and genotype VII, poultry and ostrich NDVs (Yin et al., 2011). Moreover, Chong et al. (2010) and Han et al. (2008) have reported recombinants of NDVs that involved multiple genotypes. The results of the current study present the first natural recombination event detected between lineage 2 and 5 following isolation and sequencing of virus from an ND outbreak in chicken.

Vaccination against NDV has been described as the most effective prevention strategy seconded to strict bio-security

TABLE 8 | Estimates of the evolutionary distance of NP, P, F, HN, M, L amino acid and complete genome nucleotides sequences between Malaysian NDV isolates and each lineage.

Gene	NDV isolate	Lineage 1	Lineage 2	Lineage 3a	Lineage 3b	Lineage 3c	Lineage 3d	Lineage 4	Lineage 5
NP	MB128/4	5.51	7.72	5.68	4.61	4.07	4.03	4.31	2.96
	MB076/5	5.51	7.72	5.68	4.61	4.07	4.03	4.31	2.96
	IBS002/11	6.68	9.28	7.21	5.54	4.5	4.03	5.21	3.22
	IBS005/11	6.9	9.51	7.43	5.76	4.71	4.25	5.34	3.43
	IBS025/13	3.67	0.65	7.21	7.65	7.10	8.02	8.76	8.08
P	MB128/4	15.65	17.17	14.43	13.16	13.16	12.41	13.01	9.75
	MB076/5	15.65	17.17	14.43	13.16	13.16	12.41	13.01	9.75
	IBS002/11	17.62	18.52	15.7	14.68	14.68	14.18	14.43	9.82
	IBS005/11	17.62	18.52	15.7	14.68	15.06	14.68	14.94	10.23
	IBS025/13	9.16	2.62	11.90	12.83	16.08	14.26	16.61	16.46
M	MB128/4	9.29	12.64	10.99	9.98	10.30	6.50	5.60	4.26
	MB076/5	9.29	12.64	10.99	9.98	10.30	6.50	5.60	4.26
	IBS002/11	7.69	11.26	9.89	9.89	10.30	5.40	5.27	4.04
	IBS005/11	7.42	10.99	9.62	9.62	10.03	5.13	4.73	3.49
	IBS025/13	8.46	11.81	10.71	9.62	11.13	6.50	5.60	4.42
F	MB128/4	12.38	14.59	12.33	11.53	10.59	9.03	8.65	5.66
	MB076/5	12.38	14.59	12.33	11.53	10.59	9.03	8.65	5.66
	IBS002/11	14.03	15.68	13.42	12.84	11.88	10.69	10.67	6.58
	IBS005/11	14.21	15.68	13.6	13.06	11.85	10.87	10.95	6.76
	IBS025/13	13.19	15.62	13.06	12.90	12.06	10.43	10.17	6.29
HN	MB128/4	9.95	12.26	10.68	11.79	8.41	7.53	7.64	4.54
	MB076/5	9.95	12.26	10.68	11.79	8.41	7.53	7.64	5.15
	IBS002/11	10.68	13.34	11.21	12.96	9.19	8.06	8.72	8.72
	IBS005/11	10.19	12.81	10.68	12.43	9.02	7.76	8.44	5.81
	IBS025/13	10.61	13.11	11.91	12.84	9.63	8.41	8.65	6.01
L	MB128/4	5.17	6.54	5.04	4.52	4.11	4.18	4.21	2.44
	MB076/5	5.17	6.54	5.04	4.52	4.11	4.18	4.21	2.44
	IBS002/11	5.25	6.55	5.4	4.93	4.31	4.18	4.28	2.56
	IBS005/11	5.2	6.51	5.27	4.77	4.18	4.04	4.16	2.39
	IBS025/13	5.12	6.56	5.13	4.65	4.15	4.25	4.41	2.55
Complete	MB128/4	17.93	19.79	16.14	15.21	12.88	13.09	11.98	7.87
	MB076/5	17.90	19.83	16.10	15.17	12.84	13.04	11.92	7.82
	IBS002/11	18.85	20.56	16.78	16.27	14.28	14.17	13.45	8.62
	IBS005/11	18.82	20.37	16.72	16.16	14.26	14.12	13.35	8.49
	IBS025/13	16.73	14.62	16.11	15.92	15.03	14.99	14.61	10.09

Underlined values show the minimum evolutionary distances between each lineage and Malaysian NDV isolates. Using MEGA6 program, the data was generated and values indicates% nucleotide and amino acids distance respectively.

measures in modern poultry farming where live and killed vaccines are widely used throughout various countries in the world. Despite intensive vaccination, the occurrence of NDV outbreaks in endemic countries in Asia, Africa, and Central America is puzzling (Czeglédi et al., 2006). High-density rearing in modern poultry farming, which enhances close animal-to-animal contact, favoring transmission of virulent viruses over milder forms and selective immune pressure exerted by improper vaccination, can affect the evolution of circulating virulent viruses (Miller et al., 2010; Zenglei and Liu, 2015). Additionally, improper vaccination strategy, immune suppression and presence of variant NDV strains have been implicated as the main underlying factors of poor vaccine efficacy against ND (Okoye and Shoyinka, 1983; Yu et al., 2001; Cho et al., 2007; Dortmans et al., 2012; Perozo et al., 2012).

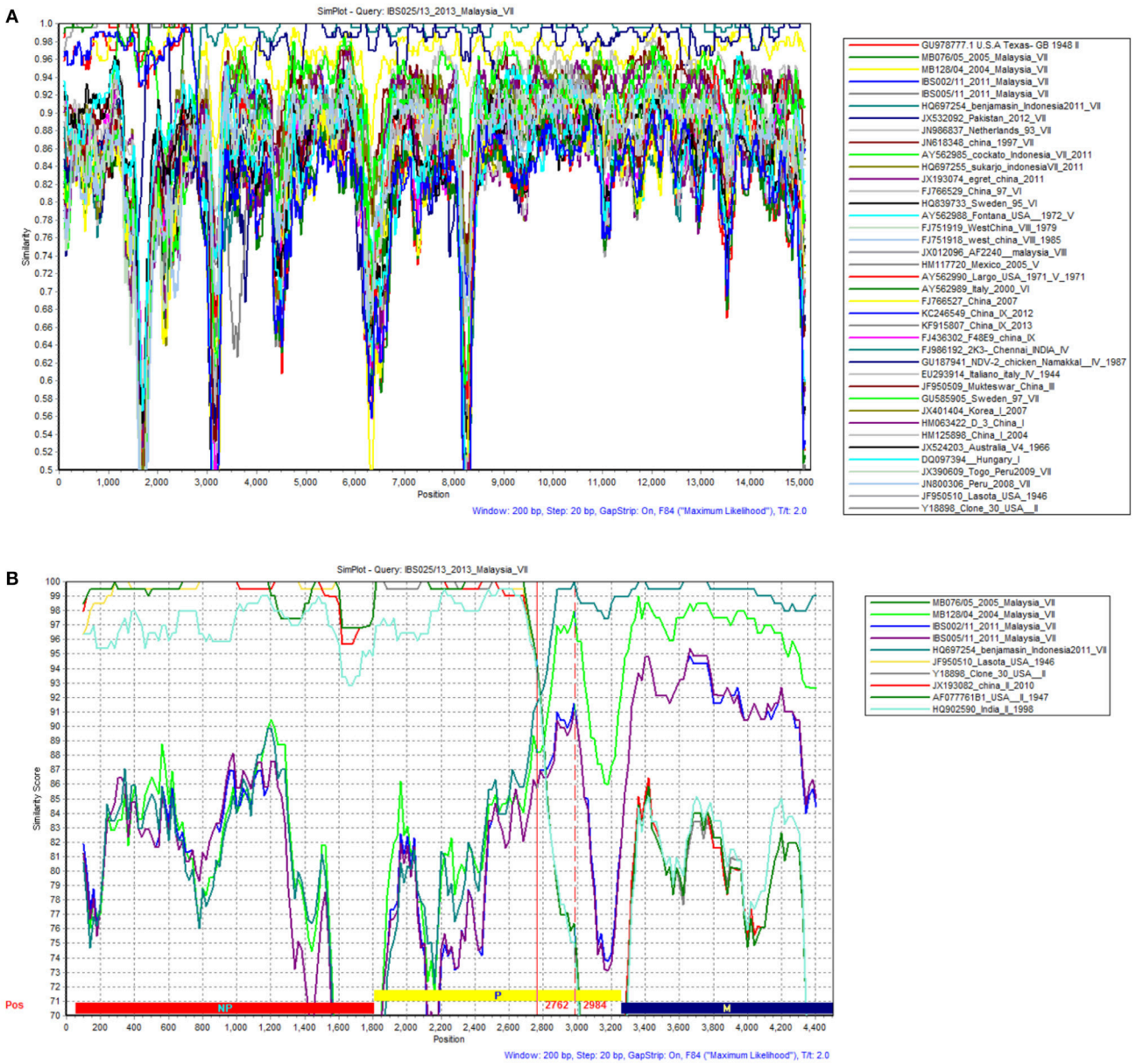


FIGURE 3 | Results from SimPlot Analysis of IBS025/13. The evidence for recombination of strain IBS025/13. **(A)** The similarity of the complete genome of IBS025/13 from SimPlot analysis with 38 representative complete genomes belonged to APMV-1. Standard similarity plot constructed using all sites of the complete genome within a sliding window of 200 bp wide centered on the position plotted with a step size between plots of 20 bp. The y-axis gives the percentage of identity. **(B)** The similarity of the first 1/3 genome representing NP, P, and M gene of IBS025/13 from SimPlot analysis with most similar nucleotide sequences belong to APMV-1. Standard similarity plot constructed using all sites of the first 1/3 genome with a window size of 200 bp and a step size of 20 bp. The y-axis gives the percentage of identity. Red vertical lines indicate the possible breakpoints.

Based on bioinformatics analysis, it has been reported that vaccination with live attenuated viruses altered the evolution of NDV by sustaining a large effective population size of a vaccine-related genotype, allowing for co-infection and recombination of vaccine and wild-type strains (Chong et al., 2010). In that study, recombination events in NP gene of Cockatoo/14698 Indonesia (AY562985) and F gene of Layer/SRZ03 China (EU167540) of genotype VII isolate with genotype I isolates were identified by maximum likelihood trees and RDP3 program (Chong et al., 2010). The current study provides the evidence of recombination in NDV between vaccine lineage and the circulating lineage 5 following an isolation and sequencing. Since live attenuated vaccines are being carried out in most of the NDV-endemic countries, the likelihood of recombination between vaccine and field strains is relatively high. The magnitude of occurrence and underlying molecular mechanisms of natural recombination

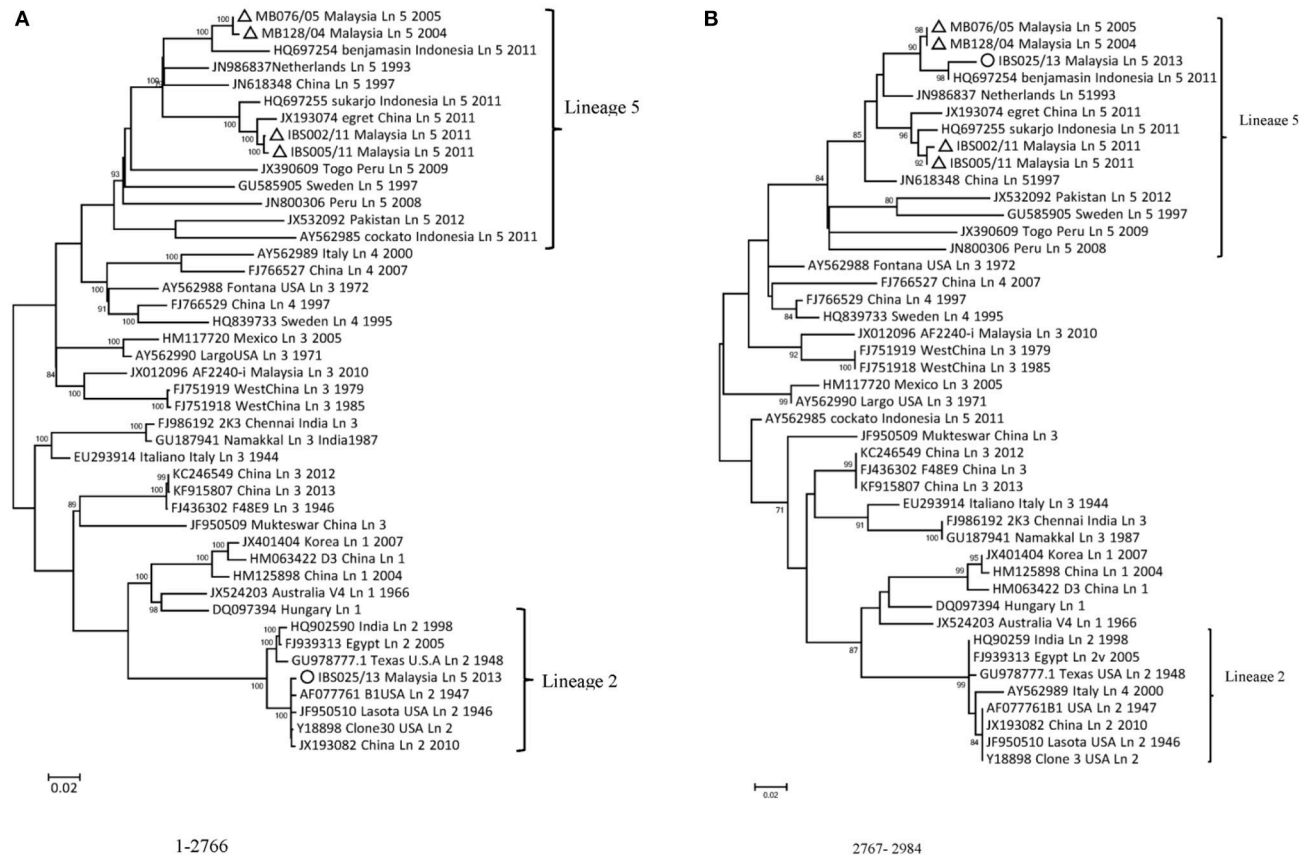


FIGURE 4 | Neighbor-joining phylogenies inferred for the two regions (A,B) delimited by the breakpoints. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (only N70% is shown). The tree is drawn to a scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The Maximum Composite Likelihood model as described in Methods, and are in the units of the number of base substitutions per site was used to compute the evolutionary distances. Codon positions included were 1st + 2nd + 3rd + non-coding. All the positions that contained gaps and missing data were removed completely from the dataset (complete deletion option). Phylogenetic analyses were conducted in MEGA6. Malaysian isolates presented in this study are marked with a triangle (Δ) and the putative mosaic is indicated with a circle (○).

in NDV is not well documented. Many research groups have documented natural recombination events for NDV (Han et al., 2008; Qin et al., 2008; Chong et al., 2010; Zhang et al., 2010; Yin et al., 2011). However, it has also been suggested that the recombination between different NDV strains is a rare event (Afonso, 2008) and that apparent genetic recombination in NDV may be an artifact (artificial recombination) (Song et al., 2011). Nevertheless, evidence of recombination between wild type and vaccine strains of NDV is relatively under-reported compared to other RNA viruses including avian influenza viruses (Webster et al., 1974), infectious bursal disease virus (Hon et al., 2008) and infectious bronchitis virus (Cavanagh et al., 1992).

Taken together, this study has described the application of NGS-based technology for complete genome sequencing of genotype VII NDV isolates. The protocol was able to generate consensus-level full genome sequence of five virulent NDVs obtained from outbreaks during 2004–2013 in Malaysia. The in-depth studies conducted on the different genotype VII NDV demonstrated the increase in the evolutionary distance of F and HN proteins of circulating lineage 5 NDVs against lineage 2. Furthermore, this study provides the evidence of recombination between the vaccine lineage and circulating virus, which warrants the importance of continuous investigations on the genome-wide study of NDV diversity.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: AO, DS, ST, and SY. Performed the experiments: DS and KM. Analyzed the data: DS, AO, ST, and MM. Contributed reagents/materials/analysis tools: AO and AI. Wrote the manuscript: DS. Critical revision: AO and MM.

FUNDING

This work was supported by an Institute of Bioscience, Higher Institution Centre of Excellence grant (IBS HICoE 6369101) from the Ministry of Education, Government of Malaysia,

Universiti Putra Malaysia Grant No. P-IPB/2013/9425700, and Biotechnology and Biological Sciences Research Council (BBSRC) through Institute Strategic Program Grant (BB/J004448/1).

ACKNOWLEDGMENTS

We thank the entire technical staff of the Laboratory of Vaccine and Immunotherapeutic for their excellent assistant.

REFERENCES

Afonso, C. L. (2008). Not so fast on recombination analysis of Newcastle disease virus. *J. Virol.* 82, 9303–9303. doi: 10.1128/JVI.01231-08

Aldous, E. W., Mynn, J. K., Banks, J., and Alexander, D. J. (2003). A molecular epidemiological study of avian paramyxovirus type 1 (Newcastle disease virus) isolates by phylogenetic analysis of a partial nucleotide sequence of the fusion protein gene. *Avian Pathol.* 32, 239–257. doi: 10.1080/0307945031000097 83

Alexander, D. J. (1997). "Newcastle disease and avian paramyxovirus infections," in *Diseases of Poultry*, eds B. W. Calnek, H. J. Barnes, C. W. Beard, and L. R. McDougald (Ames, IA: Iowa State University Press), 541–569.

Ballagi-Pordány, A., Wehmann, E., Herczeg, J., Belák, S., and Lomniczi, B. (1996). Identification and grouping of Newcastle disease virus strains by restriction site analysis of a region from the F gene. *Arch. Virol.* 141, 243–261. doi: 10.1007/BF01718397

Berhanu, A., Ideris, A., Omar, A. R., and Bejo, M. H. (2010). Molecular characterization of partial fusion gene and C-terminus extension length of haemagglutinin-neuraminidase gene of recently isolated Newcastle disease virus isolates in Malaysia. *Virol. J.* 7:183. doi: 10.1186/1743-422X-7-183

Cavanagh, D., Davis, P. J., and Cook, J. K. (1992). Infectious bronchitis virus: evidence for recombination within the Massachusetts serotype. *Avian Pathol.* 21, 401–408. doi: 10.1080/03079459208418858

Chambers, P., Millar, N. S., and Emmerson, P. T. (1986). Nucleotide sequence of the gene encoding the fusion glycoprotein of Newcastle disease virus. *J. Gen. Virol.* 67, 2685–2694. doi: 10.1099/0022-1317-67-12-2685

Cho, S. H., Kim, S. J., and Kwon, H. J. (2007). Genomic sequence of an antigenic variant Newcastle disease virus isolated in Korea. *Virus Genes* 35, 293–302. doi: 10.1007/s11262-007-0078-z

Chong, Y. L., Padhi, A., Hudson, P. J., and Poss, M. (2010). The effect of vaccination on the evolution and population dynamics of avian paramyxovirus-1. *PLoS Pathog.* 6:e1000872. doi: 10.1371/journal.ppat.1000872

Collins, M. S., Bashiruddin, J. B., and Alexander, D. J. (1993). Deduced amino acid sequences at the fusion protein cleavage site of Newcastle disease viruses showing variation in antigenicity and pathogenicity. *Arch. Virol.* 128, 363–370. doi: 10.1007/BF01309446

Commission of the European Communities (CEC) (1992). Council Directive 92/66/EC of 14 July 1992 introducing community measures for the control of Newcastle disease. *Office J. Eur. Comm.* L260, 1–20.

Czeglédi, A., Ujvári, D., Somogyi, E., Wehmann, E., Werner, O., and Lomniczi, B. (2006). Third genome size category of avian paramyxovirus serotype 1 (Newcastle disease virus) and evolutionary implications. *Virus Res.* 120, 36–48. doi: 10.1016/j.virusres.2005.11.009

de Leeuw, O., and Peeters, B. (1999). Complete nucleotide sequence of Newcastle disease virus: evidence for the existence of a new genus within the subfamily Paramyxovirinae. *J. Gen. Virol.* 80, 131–136. doi: 10.1099/0022-1317-80-1-131

Diel, D. G., da Silva, L. H., Liu, H., Wang, Z., Miller, P. J., and Afonso, C. L. (2012). Genetic diversity of avian paramyxovirus type 1: proposal for a unified nomenclature and classification system of Newcastle disease virus genotypes. *Infect. Genet. Evol.* 12, 1770–1779. doi: 10.1016/j.meegid.2012.07.012

Dortmans, J. C., Koch, G., Rottier, P. J., and Peeters, B. P. (2011). Virulence of Newcastle disease virus: what is known so far? *Vet. Res.* 42:122. doi: 10.1186/1297-9716-42-122

Dortmans, J. C. F. M., Peeters, B. P. H., and Koch, G. (2012). Newcastle disease virus outbreaks: vaccine mismatch or inadequate application? *Vet. Microbiol.* 160, 17–22. doi: 10.1016/j.vetmic.2012.05.003

European Union (EU) (2010). Directive 2010/63/EU of The European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Office J. Eur. Comm.* L276, 33–79. Available online at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:en:PDF

Han, G. Z., He, C. Q., Ding, N. Z., and Ma, L. Y. (2008). Identification of a natural multi-recombinant of Newcastle disease virus. *Virology* 371, 54–60. doi: 10.1016/j.virol.2007.09.038

Hon, C. C., Lam, T. T., Yip, C. W., Wong, R. T., Shi, M., Jiang, J., et al. (2008). Phylogenetic evidence for homologous recombination within the family Birnaviridae. *J. Gen. Virol.* 89, 3156–3164. doi: 10.1099/vir.0.2008/004101-0

Khan, T. A., Rue, C. A., Rehmani, S. F., Ahmed, A., Wasilenko, J. L., Miller, P. J., et al. (2010). Phylogenetic and biological characterization of Newcastle disease virus isolates from Pakistan. *J. Clin. Microbiol.* 48, 1892–1894. doi: 10.1128/JCM.00148-10

Kim, L. M., King, D. J., Curry, P. E., Suarez, D. L., Swayne, D. E., Stallknecht, D. E., et al. (2007a). Phylogenetic diversity among low-virulence Newcastle disease viruses from waterfowl and shorebirds and comparison of genotype distributions to those of poultry-origin isolates. *J. Virol.* 81, 12641–12653. doi: 10.1128/JVI.00843-07

Kim, L. M., King, D. J., Suarez, D. L., Wong, C. W., and Afonso, C. L. (2007b). Characterization of class I Newcastle disease virus isolates from Hong Kong live bird markets and detection using real-time reverse transcription-PCR. *J. Clin. Microbiol.* 45, 1310–1314. doi: 10.1128/JCM.02594-06

Leary, S., Underwood, W., Anthony, R., Cartner, S., Corey, D., Grandin, T., et al. (2013). *AVMA Guidelines for the Euthanasia of Animals.* Schaumburg, IL: American Veterinary Medical Association.

Locke, D. P., Sellers, H. S., Crawford, J. M., Schultz-Cherry, S., King, D. J., Meinersmann, R. J., et al. (2000). Newcastle disease virus phosphoprotein gene analysis and transcriptional editing in avian cells. *Virus Res.* 69, 55–68. doi: 10.1016/S0168-1702(00)00175-1

Maminiaina, O. F., Gil, P., Briand, F. X., Albina, E., Keita, D., Andriamanivo, H. R., et al. (2010). Newcastle disease virus in madagascar: identification of an original genotype possibly deriving from a died out ancestor of genotype IV. *PLoS ONE* 5:e13987. doi: 10.1371/journal.pone.0013987

Martin, D. P., Williamson, C., and Posada, D. (2005). RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* 21, 260–262. doi: 10.1093/bioinformatics/bth490

Miller, P. J., Decanini, E. L., and Afonso, C. L. (2010). Newcastle disease: evolution of genotypes and the related diagnostic challenges. *Infect. Genet. Evol.* 10, 26–35. doi: 10.1016/j.meegid.2009.09.012

Munir, M., Cortey, M., Abbas, M., Qureshi, Z. U., Afzal, F., Shabbir, M. Z., et al. (2012). Biological characterization and phylogenetic analysis of a novel genetic group of Newcastle disease virus isolated from outbreaks in commercial poultry and from backyard poultry flocks in Pakistan. *Infect. Genet. Evol.* 12, 1010–1019. doi: 10.1016/j.meegid.2012.02.015

Nagai, Y., Klenk, H. D., and Rott, R. (1976). Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. *Virology* 72, 494–508.

Okoye, J. O., and Shoyinka, S. V. (1983). Newcastle disease in a vaccinated flock which had experienced subclinical infectious bursal disease. *Trop. Anim. Health. Prod.* 15, 221–225. doi: 10.1007/BF02242062

Panda, A., Huang, Z., Elankumaran, S., Rockemann, D. D., and Samal, S. K. (2004). Role of fusion protein cleavage site in the virulence of Newcastle disease virus. *Microb. Pathog.* 36, 1–10. doi: 10.1016/j.micpath.2003.07.003

Perozo, F., Marcano, R., and Afonso, C. L. (2012). Biological and phylogenetic characterization of a genotype VII Newcastle disease virus from Venezuela: efficacy of field vaccination. *J. Clin. Microbiol.* 50, 1204–1208. doi: 10.1128/JCM.06506-11

Qin, Z., Sun, L., Ma, B., Cui, Z., Zhu, Y., Kitamura, Y., et al. (2008). F gene recombination between genotype II and VII Newcastle disease virus. *Virus Res.* 131, 299–303. doi: 10.1016/j.virusres.2007.10.001

Ray, S. C. (2003). *SimPlot for Windows (version 3.5.1).* Baltimore, MD. Available online at: http://sray.med.som.jhmi.edu/SCRoftware/ (Accessed June 2015).

Read, A. F., Baigent, S. J., Powers, C., Kgosana, L. B., Blackwell, L., Smith, L. P., et al. (2015). Imperfect vaccination can enhance the transmission of highly virulent pathogens. *PLoS Biol.* 13:e1002198. doi: 10.1371/journal.pbio.1002198

Roohani, K., Tan, S. W, Yeap, S. K., Ideris, A., Bejo, M. H., and Omar, A. R. (2015). Characterization of the genotype VII Newcastle disease virus (NDV) isolated from NDV vaccinated chickens and the efficacy of LaSota and recombinant genotype VII vaccines against challenge with velogenic NDV. *J. Vet. Sci.* 16, 447–457. doi: 10.4142/jvs.2015.16.4.447

Samuel, A., Nayak, B., Paldurai, A., Xiao, S., Aplogan, G. L., Awoume, K. A., et al. (2013). Phylogenetic and pathotypic characterization of Newcastle disease viruses circulating in West Africa and efficacy of a current vaccine. *J. Clin. Microbiol.* 51, 771–781. doi: 10.1128/JCM.02750-12

Snoeck, C. J., Ducatez, M. F, Owoade, A. A, Faleke, O. O, Alkali, B. R., Tahita, M. C., et al. (2009). Newcastle disease virus in West Africa: new virulent strains identified in non-commercial farms. *Arch. Virol.* 154, 47–54. doi: 10.1007/s00705-008-0269-5

Snoeck, C. J., Owoade, A. A., Couacy-Hymann, E., Alkali, B. R., Okwen, M. P., Adeyanju, A. T., et al. (2013). High genetic diversity of Newcastle disease virus in poultry in west and central Africa: cocirculation of genotype XIV and newly defined genotypes XVII and XVIII. *J. Clin. Microbiol.* 51, 2250–2260. doi: 10.1128/JCM.00684-13

Song, Q., Cao, Y., Li, Q., Gu, M., Zhong, L., Hu, S., et al. (2011). Artificial recombination may influence the evolutionary analysis of Newcastle disease virus. *J. Virol.* 85, 10409–10414. doi: 10.1128/JVI.00544-11

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tan, S. W., Ideris, A., Omar, A. R., Yusoff, K., and Hair-Bejo, M. (2010). Sequence and phylogenetic analysis of Newcastle disease virus genotypes isolated in Malaysia between 2004 and 2005. *Arch. Virol.* 155, 63–70. doi: 10.1007/s00705-009-0540-4

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.

Wang, K. C, Chen, G. Q., Jiang, W. M., Liu, S., Hou, G. Y., Yu, J. M., et al. (2013). Complete genome sequence of a hemagglutination-negative avian paramyxovirus type 4 isolated from China. *Genome Announc.* 1:e00045-13. doi: 10.1128/genomeA.00045-13

Webster, R. G., Isachenko, V. A., and Carter, M. (1974). A new avian influenza virus from feral birds in the USSR: recombination in nature? *Bull World Health Organ.* 51, 325–332.

World Animal Health Information Database Interface (WAHID) (2015). Available online at: http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/statusdetail (Accessed July 2015).

World Organisation for Animal Health (OIE) (2012). "Newcastle disease," in *OIE Manual of Standards for Diagnostic Tests and Vaccines for Terrestrial Animals,* eds S. Linnane and S. Edwards (Paris: World Organisation for Animal Health), 576–589.

Xiao, S., Nayak, B., Samuel, A., Paldurai, A., Kanabagattebasavarajappa, M., Prajitno, T. Y., et al. (2012). Generation by reverse genetics of an effective, stable, live-attenuated Newcastle disease virus vaccine based on a currently circulating, highly virulent Indonesian strain. *PLoS ONE* 7:e52751. doi: 10.1371/journal.pone.0052751

Yin, Y., Cortey, M., Zhang, Y., Cui, S., Dolz, R., Wang, J., et al. (2011). Molecular characterization of Newcastle disease viruses in Ostriches (*Struthio camelus* L.): further evidences of recombination within avian paramyxovirus type 1. *Vet. Microbiol.* 149, 324–329. doi: 10.1016/j.vetmic.2010.12.004

Yu, L., Wang, Z., Jiang, Y., Chang, L., and Kwang, J. (2001). Characterization of newly emerging Newcastle disease virus isolates from the People's Republic of China and Taiwan. *J. Clin. Microbiol.* 39, 3512–3519. doi: 10.1128/JCM.39.10.3512-3519.2001

Yusoff, K., and Tan, W. S. (2001). Newcastle disease virus: macromolecules and opportunities. *Avian Pathol.* 30, 439–455. doi: 10.1080/03079450120078626

Zenglei, H., and Liu, X. (2015). Immune response-induced NDV pathology. *Br. J. Virol.* 2, 25–27. Available online at: http://smithandfranklin.com/current-issues/NDV-Induced-Immune-Pathology-in-Chickens/6/15/93/html

Zhang, R., Wang, X., Su, J., Zhao, J., and Zhang, G. (2010). Isolation and analysis of two naturally-occurring multi-recombination Newcastle disease viruses in China. *Virus Res.* 151, 45–53. doi: 10.1016/j.virusres.2010.03.015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Satharasinghe, Murulitharan, Tan, Yeap, Munir, Ideris and Omar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 February 2016
doi: 10.3389/fmicb.2016.00119

Phylogenetic and Pathotypic Characterization of Newcastle Disease Viruses Circulating in South China and Transmission in Different Birds

Yinfeng Kang[1,2†], Bin Xiang[1,2†], Runyu Yuan[1,2,3†], Xiaqiong Zhao[1,2], Minsha Feng[1,2], Pei Gao[1,2], Yanling Li[1,2], Yulian Li[1,2], Zhangyong Ning[1] and Tao Ren[1,2]*

[1] Key Laboratory of Animal Vaccine Development, College of Veterinary Medicine, South China Agricultural University, Guangzhou, China, [2] Key Laboratory of Zoonosis Prevention and Control of Guangdong Province, Guangzhou, China, [3] Key Laboratory for Repository and Application of Pathogenic Microbiology, Research Center for Pathogens Detection Technology of Emerging Infectious Diseases, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China

OPEN ACCESS

Edited by:
Akio Adachi, Tokushima University Graduate School, Japan

Reviewed by:
Takashi Irie, Hiroshima University, Japan
Masato Tsurudome, Mie University Graduate School of Medicine, Japan

***Correspondence:**
Tao Ren
rentao6868@126.com

[†] *These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *11 November 2015*
Accepted: *22 January 2016*
Published: *09 February 2016*

Citation:
Kang Y, Xiang B, Yuan R, Zhao X, Feng M, Gao P, Li Y, Li Y, Ning Z and Ren T (2016) Phylogenetic and Pathotypic Characterization of Newcastle Disease Viruses Circulating in South China and Transmission in Different Birds. Front. Microbiol. 7:119. doi: 10.3389/fmicb.2016.00119

Although Newcastle disease virus (NDV) with high pathogenicity has frequently been isolated in poultry in China since 1948, the mode of its transmission among avian species remains largely unknown. Given that various wild bird species have been implicated as sources of transmission, in this study we genotypically and pathotypically characterized 23 NDV isolates collected from chickens, ducks, and pigeons in live bird markets (LBMs) in South China as part of an H7N9 surveillance program during December 2013–February 2014. To simulate the natural transmission of different kinds of animals in LBMs, we selected three representative NDVs—namely, GM, YF18, and GZ289—isolated from different birds to evaluate the pathogenicity and transmission of the indicated viruses in chickens, ducks, and pigeons. Furthermore, to investigate the replication and shedding of NDV in poultry, we inoculated the chickens, ducks, and pigeons with 10^6 EID_{50} of each virus via intraocular and intranasal routes. Eight hour after infection, the naïve contact groups were housed with those inoculated with each of the viruses as a means to monitor contact transmission. Our results indicated that genetically diverse viruses circulate in LBMs in South China's Guangdong Province and that NDV from different birds have different tissue tropisms and host ranges when transmitted in different birds. We therefore propose the continuous epidemiological surveillance of LBMs to support the prevention of the spread of these viruses in different birds, especially chickens, and highlight the need for studies of the virus–host relationship.

Keywords: newcastle disease virus (NDV), phylogenetic analysis, pathogenicity, transmission, South China

INTRODUCTION

Due to its high morbidity and mortality, Newcastle disease (ND) is considered to be the most significant and widespread infectious disease in commercial poultry—one that can cause severe economic losses in poultry, particularly chickens, and affect a range of other domestic species, including ducks and pigeons (Sinkovics and Horvath, 2000; OIE, 2012). In chickens, the clinical

manifestation of NDV strains varies significantly due to the degree of strain virulence and host susceptibility (Alexander, 2000). NDV strains are categorized as highly virulent (velogenic), moderately virulent (mesogenic), or avirulent (lentogenic), all according to pathogenicity for chickens gauged by the intracerebral pathogenicity index (ICPI) and mean death time (MDT; OIE, 2012).

Historically, NDV strains have been divided into two major divisions—class I and class II—which contribute to the extensive genetic diversity among poultry worldwide, with class I being further divided into nine genotypes distributed worldwide in waterfowls and class II comprising 18 (I–XVIII) genotypes when the sequences are isolated over time (Ballagi-Pordány et al., 1996; Liu et al., 2003; Kim et al., 2007; Miller et al., 2010; Diel et al., 2012; OIE, 2012; Snoeck et al., 2013). With the exception of 9a5b, class I strains are avirulent in chickens and have historically been recovered from aquatic birds (Shengqing et al., 2002; Kim et al., 2007; Diel et al., 2012). Genotypes I, II, V, VI, VII, and IX of NDV have often been isolated in South China (e.g., Guangdong Province), although genotypes V, VI, and VII of class II strains are the predominant genotypes there and contain only virulent viruses (Jinding et al., 2005; Cai et al., 2011; Kang et al., 2014; Wang et al., 2015). In 1948, genotype IX emerged as a unique group including the first virulent NDV strains (F48E9) in South China, and members of the genotype are occasionally isolated from a wild variety of bird species (Liu et al., 2003). To date, many newly emergent strains isolated from a wide range of birds contribute to increasing the global burden of NDV and cause enormous losses in the poultry industry (Liu et al., 2003; Jinding et al., 2005; Kim et al., 2007; Kang et al., 2014). In 1981, genotype VI of NDV strains was first isolated from pigeons and was repeatedly isolated in China until 1985, when genotype VII became more epidemic and posed a constant threat to domestic poultry (Mase et al., 2002; Liu et al., 2003; Cai et al., 2011). Genotype VII of NDV strains have continued to circulate in poultry throughout South China and are now considered to be enzootic as they spread around the globe.

NDV has a wide range of hosts, as more than 250 bird species have been found to be susceptible by natural or experimental infections, although wild waterfowl and shorebirds are regarded to be the reservoir of the virus in nature (Kaleta and Baldauf, 1988). Chickens, pigeons, and ducks are most commonly susceptible to infection with the same NDV strain, though they exhibit different susceptibility (Erickson et al., 1980; Cattoli et al., 2011; Smietanka et al., 2014). As the most susceptible poultry, chickens are subject to high morbidity and mortality if infected with virulent NDV (Bogoyavlenskiy et al., 2005; Kang et al., 2014). Aquatic birds, including ducks, as natural reservoirs of NDV show nearly no obvious clinical symptoms when infected with NDV strains, even those virulent in chickens (Zhang et al., 2011; Kang et al., 2014). However, since the first large-scale outbreak of ND in geese in South China in 1997, duck-origin NDV strains have exhibited high virulence in waterfowl (Jinding et al., 2005; Zhang et al., 2011).

Some type 1 pigeon paramyxoviruses (PPMV-1) behave similarly to non-virulent viruses according to ICPI tests in 1-day-old chickens; whereas the strains are highly pathogenic to pigeons, the serial passaging of PPMV-1 in chickens results in increased virulence (Kommers et al., 2003; Dortmans et al., 2011). As a result, upon the natural transmission from pigeons to chickens, PPMV-1 strains may evolve into virulent viruses and induce major outbreaks (Dortmans et al., 2011). In China, the implementation of extensive vaccination procedures among commercial poultry farms and the culling of infected poultry have reduced the number of epizootic outbreaks of ND since the 1980s (Liu et al., 2003; Kang et al., 2014). However, genotypes V, VI, VII, and IX of NDV continue to circulate and frequently cause outbreaks in China (Cai et al., 2011; Qiu et al., 2011; Zhang et al., 2011; Kang et al., 2014; Wang et al., 2015). Nevertheless, very little information is available regarding the epidemiology, evolutionary trends, and transmission of the virus circulating in South China.

Previous studies have shown that PPMV-1 can be transmitted from infected pigeons to chickens placed in naïve contact (Alexander and Parsons, 1984; de Oliveira Torres Carrasco et al., 2008). At the same time, duck-origin NDV isolate can infect chickens and ducks and be transmitted to naïve contact chickens and ducks (Kang et al., 2014). However, little information is available regarding the pathogenesis and transmission of NDV among chickens, ducks, and pigeons. In this study, we therefore investigated the presence of NDV in chickens, ducks, and pigeons in live bird markets (LBMs) in South China and selected three viruses isolated from the different birds to better evaluate the pathogenicity and transmission of the NDV among various types of bird.

MATERIALS AND METHODS

Ethics Statement

The present study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the Ministry of Science and Technology of the People's Republic of China. All animal experiments were performed in animal biosafety level 3 facilities and were conducted under the guidance of South China Agricultural University's Institutional Animal Care and Use Committee and the Association for Assessment and Accreditation of Laboratory Animal Care International accredited facility. The protocol was reviewed and approved by the Committee on the Ethics of Animal Experiments of the animal biosafety level 3 Committee of South China Agricultural University, the approve ID is SCXK (Guangdong) 2013-0019.

Virus Isolation and Biological Characterization

Isolates were collected in LBMs in South China's Guangdong Province as part of an avian influenza (H7N9) surveillance program during December 2013–February 2014 by the Key Laboratory of Zoonosis Prevention and Control of Guangdong, China. Oropharyngeal and cloacal swab specimens were collected from commercial chickens, ducks, and pigeons from LBMs, also in Guangdong Province. A total of 214 swab samples were collected in 1.5 mL centrifuge tubes containing 1.0 mL

transport medium (50% glycerol in phosphate-buffered saline [PBS]) with antibiotics (penicillin, 2000 U/mL; amphotericin B, 2000 U/mL; streptomycin, 2 mg/mL) and shipped to South China Agricultural University. The oropharyngeal and cloacal swab samples were inoculated in 10-days-old specific-pathogen-free (SPF) embryonated chicken eggs, as previously described (Kang et al., 2015). A hemagglutination (HA)-inhibition (HI) test was conducted using four HA units of the isolates and virus purification performed, as described by Office International Des Epizooties (OIE; OIE, 2012). A total of 23 NDV strains were isolated (**Table 1**). The MDT in 9-days-old SPF embryonated chicken eggs and ICPI tests in 1-day-old SPF chickens were performed according to OIE's (2012) standard procedure. Isolates were titrated in 10-days-old SPF-embryonated chicken eggs and stored at −80°C for further characterization. The evaluation of 50% egg infective doses (EID_{50}) was calculated by using the Reed–Muench method (Reed and Muench, 1938).

Animals

To test pathogenicity and reproduce the natural conditions of NDV transmission in chickens, ducks, and pigeons, we selected three strains representative of and circulating in South China: genotype VII Chicken/Guangdong/GM/2014 (GM), genotype IX Duck/Guangdong/YF18/2014 (YF18), and genotype VI Pigeon/Guangdong/GZ289/2014 (GZ289), each a predominant chicken-origin, duck-origin, and pigeon-origin genotype, respectively. A total of 27 6-weeks-old SPF white Leghorn chickens were supplied by Guangdong Dahuanong and housed in isolator cages under negative pressure with food and water provided *ad libitum*. A total of 27 pigeons (15-weeks-old *Columba livia* rock pigeons) and 27 commercially available domestic ducks (2-weeks-old Peking ducks) were purchased from a pigeon farm in Gaoming and a duck farm in Yunfu, respectively, and housed in isolators. All pigeons and ducks were confirmed to be serologically negative for ND by HI assays during a 1-day period before experimentation.

Genetic and Phylogenetic Analyses

We previously determined the complete genome sequence of the GM strain from LBMs in Guangdong Province preserved in our laboratory (Wang et al., 2014). The complete genome of the other two viruses YF18 and GZ289 used in this study were

TABLE 1 | Epidemiological, genetic, and biological characteristics description of 23 NDVs isolated from South China.

NDV isolate[a]	Host	Location (City)	Class	Genotype	MDT[b](h)	ICPI[c]	F-protein cleavage site[d]	Pathotype	GenBank accession no.[e]
Pigeon/Guangdong/GZ293/2014*	Pigeon	Guangzhou	Class II	VI	78	1.37	RRQKRF	Velogenic	KT381592
Chicken/Guangdong/GM5/2013 &*	Chicken	Gaoming	Class II	VII	70	1.73	RRQKRF	Velogenic	KT381593
Chicken/Guangdong/GM/2014 &*[f]	Chicken	Gaoming	Class II	VII	56	1.78	RRQKRF	Velogenic	DQ486859
Duck/Guangdong/YF21/2014 &*	Duck	Yunfu	Class II	VII	58	1.65	RRQKRF	Velogenic	KT381594
Pigeon/Guangdong/GM1/2014 &	Pigeon	Gaoming	Class II	VI	64	1.41	RRQKRF	Velogenic	KT381595
Pigeon/Guangdong/GM8/2013*	Pigeon	Gaoming	Class II	VI	87	1.29	KRQKRF	Velogenic	KT381596
Chicken/Guangdong/GZ295/2014	Chicken	Guangzhou	Class II	I	>120	0.5	RKQGRL	Lentogenic	KT381597
Duck/Guangdong/GM12/2014	Duck	Gaoming	Class II	I	>120	0.2	RKQGRL	Lentogenic	KT381598
Duck/Guangdong/GZ331/2014	Duck	Guangzhou	Class II	I	>120	0.2	RKQGHL	Lentogenic	KT381599
Duck/Guangdong/YF24/2014 &*	Duck	Yunfu	Class II	II	>120	0.2	GRQGRL	Lentogenic	KT381600
Pigeon/Guangdong/GZ287/2013 &	Pigeon	Guangzhou	Class II	VI	78	1.22	RRQKRF	Velogenic	KT381601
Pigeon/Guangdong/GZ288/2013 &	Pigeon	Guangzhou	Class II	VI	76	1.22	RRQKRF	Velogenic	KT381602
Pigeon/Guangdong/GZ290/2013	Pigeon	Guangzhou	Class II	VI	84	1.35	RRQKRF	Velogenic	KT381603
Pigeon/Guangdong/GZ292/2014	Pigeon	Guangzhou	Class II	VI	78	1.38	RRQKRF	Velogenic	KT381604
Duck/Guangdong/YF18/2014 &*	Duck	Yunfu	Class II	IX	42	1.83	RRQRRF	Velogenic	KT381605
Pigeon/Guangdong/GZ289/2014 &	Pigeon	Guangzhou	Class II	VI	58	1.72	RRQKRF	Velogenic	KT381606
Duck/Guangdong/YF821/2014 &*	Duck	Yunfu	Class I	3	>120	0	ERQERL	Lentogenic	KT381585
Duck/Guangdong/YF827/2014 &*	Duck	Yunfu	Class I	3	>120	0	ERQERL	Lentogenic	KT381586
Pigeon/Guangdong/YF1/2014	Pigeon	Yunfu	Class I	3	>120	0	ERQERL	Lentogenic	KT381587
Chicken/Guangdong/GM3/2013	Chicken	Gaoming	Class I	3	>120	0	ERQERL	Lentogenic	KT381588
Chicken/Guangdong/GM307/2014	Chicken	Gaoming	Class I	3	>120	0	ERQERL	Lentogenic	KT381589
Chicken/Guangdong/GZ316/2014	Chicken	Guangzhou	Class I	3	>120	0	ERQERL	Lentogenic	KT381590
Chicken/Guangdong/GZ379/2014	Chicken	Guangzhou	Class I	3	>120	0	ERQERL	Lentogenic	KT381591

[a] *NDV was isolated from the oropharyngeal swab samples (&) or from both oropharyngeal and cloacal swab samples from the same bird (*). NDV isolates without any symbolic notation were isolated from cloacal swabs only.*

[b] *Mean death time in 10 d-old SPF embryonated chicken eggs (hours) (<60, velogen; 60–90, mesogen; >90, lentogen).*

[c] *Intracerebral pathogenicity index in day-old chickens (<0.7, lentogen; 0.7–1.4, mesogen; 1.4–2.0, velogen).*

[d] *Amino acids 112–117.*

[e] *The Gen Bank accession numbers provided are for the 1662-bp nucleotide sequence of the NDV-F gene open reading frame of South China NDV strains.*

[f] *Published by Wang et al. (2014).*

sequenced. For the other 20 isolates, viral genomic RNA was extracted from infected allantoic fluid by using an RNA isolation kit (RNeasy Mini Kit, Qiagen, Hilden, Germany) according to the manufacturer's instructions and reverse transcribed into cDNA with a cDNA synthesis kit (SuperScript RT-PCR, Invitrogen, Carlsbad, CA, USA). PCR amplification was carried out using primers specific to the complete genomes of NDV, as described previously (Kang et al., 2014), and the NDV F gene open reading frame was amplified using primers, also as described previously (Kang et al., 2014). PCR products were purified with a DNA purification kit (Corning, NY, USA) and sequenced using a cycle sequencing kit (BigDye Terminator, Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. Nucleotide sequences were compiled and edited using Lasergene version 7.1 software (DNASTAR, Madison, WI, USA). A phylogenetic tree based on the F gene open reading frame and the complete genome of the NDVs were constructed by using the maximum likelihood method with the generalized time reversible GTR+G+I4 model, using molecular evolutionary genetics analysis software (MEGA version 5.02), in following (Tamura et al., 2011; Diel et al., 2012). The statistical significance of the tree was assessed with a bootstrap value of 1000.

Animal Experiments

To assess the pathogenicity and transmission of the three representative isolated NDV strains GM, YF18, and GZ289, in chickens, ducks, and pigeons, three groups of 6-weeks-old SPF chickens, 2-weeks-old Peking ducks, and 15-weeks-old *C. livia* rock pigeons (nine birds per group) were inoculated with 10^6 EID_{50} of the indicated virus in a volume of 200 μL via intraocular and intranasal routes. Eight hour after infection, an additional three chickens, ducks, and pigeons were inoculated intranasally with 200 μL PBS and placed in physical contact—that is, in the same cage and sharing food and water *ad libitum*—with inoculated birds in order to monitor contact infection. We euthanized three inoculated birds (dead or ill with depression, torticollis, incoordination, and tremors) from each subgroup 3 days postinoculated (DPI) to test the virus replication from various tissues from the thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen. Oropharyngeal and cloacal swabs from all infected and contacted birds were collected for the detection of viruses shedding at 3, 5, 7, 9, and 11 DPI and suspended in 1000 μL transport medium (50% glycerol in PBS) with antibiotics (penicillin, 2000 U/mL; amphotericin B, 2000 U/mL; streptomycin, 2 mg/mL) for viral detection and titration in eggs. Virus titers were calculated as described previously (Kang et al., 2014). All birds were observed daily every 8 h for illness or death during the course of the 14-days experimental period. We collected serum samples from each surviving bird for serologic testing at 14 DPI. All samples were confirmed to show seroconversion by HI test using four HA units of isolates based on standard procedures (OIE, 2012).

Nucleotide Sequence Accession Numbers

Gen Bank accession numbers of the complete F genes of the 23 South China strains described in this study are designated in **Table 1** (KT381585 to KT381606). The complete genome sequences of the GM, YF18, and GZ289 strains obtained in the present study are available from Gen Bank under the accession numbers DQ486859, KR014814, and KR014815, respectively.

RESULTS

Pathogenicity Analysis

Samples of chickens, ducks, and pigeons were taken at LBMs in Guangzhou, Yunfu, and Gaoming, China, where 42 chickens (*Gallus gallus*), 38 ducks (Peking), and 27 pigeons (*C. livia*) had oropharyngeal and cloacal swab samples taken as part of an avian influenza (H7N9) surveillance program (**Table 1**). **Table 1** presents the initial biological characterizations of 23 NDV isolates, including MDT and ICPI. The seven class I genotype III and four class II genotype I or II samples with MDTs of >120 and with ICPIs of 0–0.2 h were consistent with a lentogenic pathotype. The other 12 strains from South China had MDTs of 42–87 h and ICPIs of 1.22–1.83 and were thus considered velogenic strains (OIE, 2012). **Table 1** also lists the F protein cleavage site amino acid sequence of the 23 field isolates of NDV derived in this study, consistent with the virulence pathotype identified by MDT and ICPI tests.

Genetic and Phylogenetic Analyses of NDV Isolates

Based on the complete F genome sequences derived from different birds (**Figure 1A**), three strains were assigned to genotypes I, one to genotype II, eight to genotype VI, three to genotype VII, and one to genotype IX, for a total of 16 samples in class II. Seven samples were positive for avirulent class I strains. To further determine their molecular characteristics, we selected three strains noted in the previous section for the complete genome sequenced. The sequences were compared with 52 complete representative NDV genome sequences obtained from Gen Bank. Based on complete nucleotide sequences, phylogenetic analysis revealed that the GM, YF18, and GZ289 strains clustered into class II genotypes VII, IX, and VI, respectively (**Figure 1B**). In sum, we have shown that 23 NDV strains derived from chickens, ducks, and pigeons in LBMs were virulent or avirulent circulating in South China and have confirmed the coexistence of different genotypes of NDV in domestic poultry *in vivo*, thereby suggesting that NDV may be transmitted between types of domestic poultry and pigeons in South China.

Pathogenicity and Transmission of NDV among SPF Chickens

To evaluate tissue tropism and pathogenicity of three representative South China strains GM, YF18, and GZ289, in chickens, we inoculated each chicken with 10^6 EID_{50} of infective allantoic fluid of the indicated virus in a 200 μL volume via intraocular and intranasal routes and euthanized three chickens from each subgroup at 3 DPI, after which the remaining chickens were observed clinically for 2 weeks.

All chickens exposed to GM died within 3 DPI. All chickens inoculated with YF18 died by 5 DPI. However, none

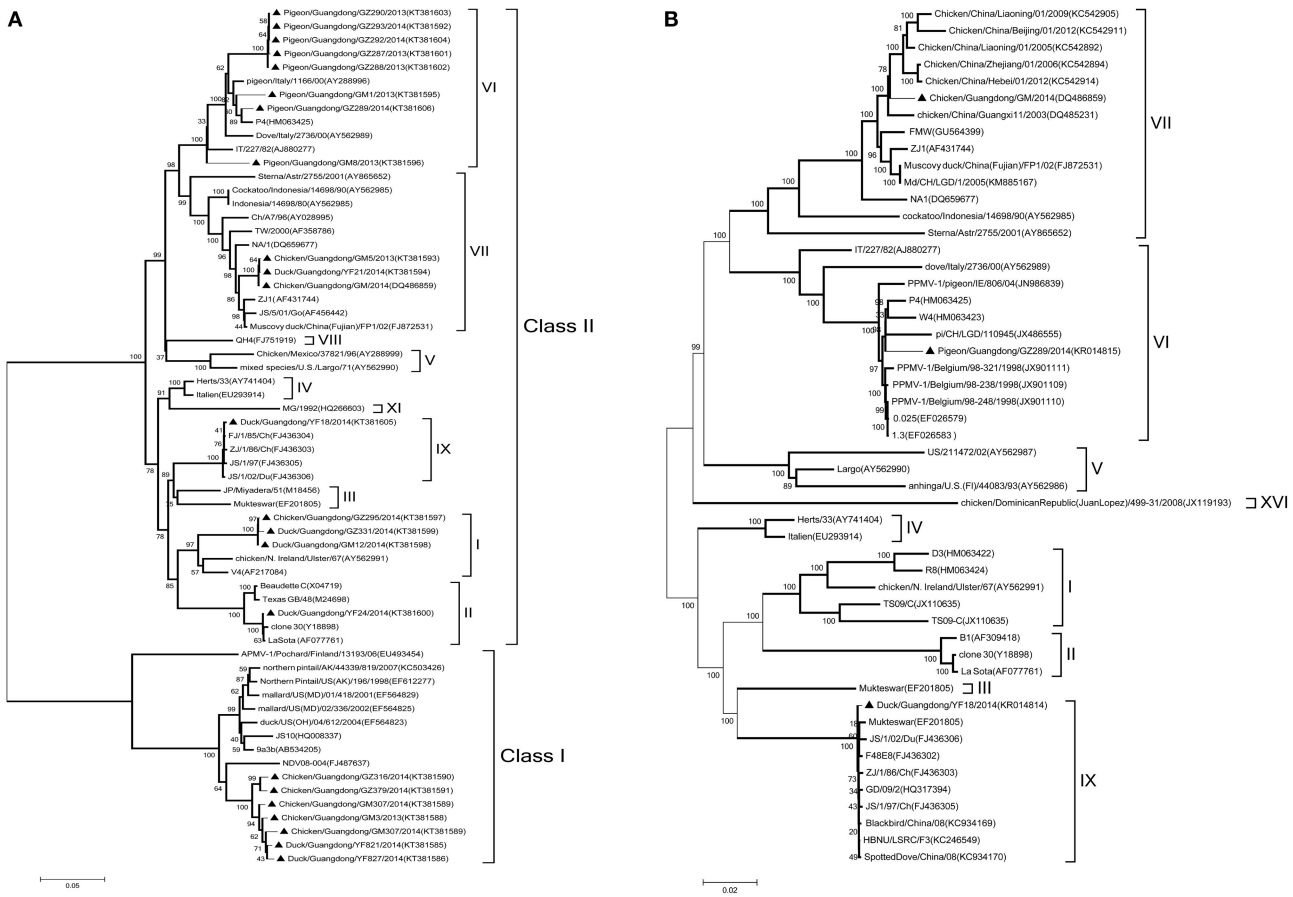


FIGURE 1 | Phylogenetic tree of NDV strains. (A) Phylogenetic analysis based on the F-gene open reading frame (1662 nt) of viruses representing NDV class I and II available in Gen Bank. **(B)** Phylogenetic analysis based on the complete genome sequence (15192 nt) of the GM, YF18, and GZ289 strains derived from different birds. Viruses highlighted with the black triangles (▲) were characterized in this study. The tree was constructed using the Maximum Likelihood method with the Generalized Time Reversible GTR+G+I4 model of software MEGA (version 5.02), with 1000 bootstrap trials to assign confidence to the groupings.

of the chickens infected GZ289 virus died. No symptom was observed in the GZ289-inoculated chickens despite their seroconversion; none of the naïve contact-group chickens seroconverted (**Table 2**). The GM, YF18, and GZ289 isolates were replicated systemically in various tissues of SPF chickens on 3 DPI, including thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen (**Table 2**). The GM, YF18, and GZ289 viruses showed remarkable replication in the lungs; with mean titers were 8.75, 8.25, and 3.42 $\log_{10}EID_{50}$, respectively. In addition, the selected three viruses furthermore replicated to the mean titer of 6.75, 6.75, 2.00 $\log_{10}EID_{50}$ in the spleen, respectively. In the GM inoculated chicken, the mean titers were from 6.50 to 8.75 $\log_{10}EID_{50}$ in the tested tissues, which were greater than those in the YF18 and GZ289-inoculated chickens (**Table 2**). In other words, the GM virus replicated more highly in the chickens.

GM, YF18, and GZ289 viruses shedding from the inoculated chickens were detected in oropharyngeal and cloacal swabs at 3, 5, 7, 9, and 11 DPI (**Table 3**). In the infected chickens, the GM virus was recovered from the oropharyngeal (4.79 $\log_{10}EID_{50}$)on 3 DPI and from the cloacal (5.29 $\log_{10}EID_{50}$) on 3 DPI (**Table 3**). The YF18 virus was shed from the oropharynx in inoculated chickens within 5 DPI (2.50–5.25 $\log_{10}EID_{50}$). But, it only could be shed from the cloaca within 3 DPI (4.92 $\log_{10}EID_{50}$). GZ289 virus shedding was detected from oropharyngeal and cloacal swabs only at 3 DPI (3.29 and 2.88 $\log_{10}EID_{50}$, respectively).

To determine whether these three viruses could be horizontally transmitted among chickens, 8 h after infection, three chickens were inoculated with 200 μL PBS via the same routes as a naïve contact group and housed with those inoculated with the GM, YF18 or GZ289 viruses. In naïve contact-group chickens, housed with inoculated GM chickens during the observed time, died within 7 DPI (**Table 2**). GM virus was recovered from the oropharyngeal swabs (4.00–4.13 $\log_{10}EID_{50}$) and from the cloacal swabs (3.75–3.92 $\log_{10}EID_{50}$) at 3 and 5 DPI (**Table 3**). The naïve contact chickens housed with YF18 died within 7 DPI (**Table 2**). The virus shedding was detected from oropharyngeal and cloacal swabs at 3, 5, and 7 DPI (3.00–4.00 $\log_{10}EID_{50}$ and 3.00–4.17 $\log_{10}EID_{50}$, respectively). The naïve contact chickens housed with GZ289 virus still survived in 14 DPI, but none seroconverted (**Table 2**). Meanwhile, GZ289 virus was not recovered from oropharyngeal

TABLE 2 | Lethality, seroconversion, and tissues tropism among chickens in an intraspecies study of NDV transmission[a].

Viruses	Manifestations of chickens				Virus replication on 3 DPI ($log_{10}EID_{50}/g$)[d] in:						
	No. D/S/total[b]		No. S.C./total[c]								
	Infected[e]	Contact[f]	Infected	Contact	Thymus	Cecal tonsils	Bursa of Fabricius	Trachea	Lung	Brain	Kidney	Spleen
GM	3/3/3	3/3/3	–	–	7.33 ± 0.14	6.92 ± 0.58	6.83 ± 0.14	5.50 ± 0.43	7.75 ± 0.50	6.13 ± 0.53	7.17 ± 0.14	6.75 ± 0.25
YF18	3/3/3	3/3/3	–	–	6.25 ± 0.50	6.83 ± 0.80	6.58 ± 0.58	5.42 ± 0.29	7.25 ± 0.90	5.50 ± 0.90	7.08 ± 0.52	6.75 ± 0.66
GZ289	0/0/3	0/0/3	3/3 (6.0)[g]	0/3	2.92 ± 0.52	2.58 ± 1.01	2.42 ± 0.58	2.33 ± 0.29	3.42 ± 0.14	2.00 ± 0.43	2.17 ± 0.58	2.00 ± 0.25

[a] *Chickens were inoculated with 10^6 EID_{50} of the indicated virus in a volume of 200 μL via the intraocular and intranasal routes; Tissues, including thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen were collected aseptically on three DPI from three chickens when infected with the indicated virus and homogenized for virus titration in eggs.*

[b] *No. D/S/total shows the number of dead (D) and sick (S) as well as the total number of chickens from the observation period. The chickens that showed clinical signs, such as depression, torticollis, incoordination and tremors, but recovered at the end of the observation were counted as sick animals.*

[c] *No. S.C./total shows the number of chickens that seroconverted out of the total number of chickens at the end of the observation period. –, all of the chickens died at the end of the observation.*

[d] *For statistical analysis, a value of 1.5 was assigned if the virus was not detected from the undiluted sample in three SPF embryonated chicken eggs (Kang et al., 2015).Virus titers are expressed as means ± standard deviation in $log_{10}EID_{50}/g$ of tissue.*

[e] *Chickens inoculated with virus.*

[f] *Three uninoculated chickens were co-housed with infected chickens as a contact group 8 h after inoculation.*

[g] *Average antibody titer of infected chickens (log_2).*

and cloacal swabs during the trial period from contact birds (**Table 3**).

Our study indicated that GM and YF18 were highly pathogenic to chickens, and could be transmitted by contact with naïve chickens, while the GZ289 virus did not replicate well in chickens, and did not spread by naïve contact (**Figure 2**).

Pathogenicity and Transmission of NDV among Domestic Ducks

To assess the tissue tropism and pathogenicity of the three viruses in ducks, we inoculated each duck with 10^6 EID_{50} of the indicated virus in 200 μL via intraocular and intranasal routes and euthanized three ducks from each subgroup at 3 DPI. The remaining ducks were observed clinically for 14 days.

All ducks in the three infected virus groups survived during the period of infection (**Table 4**). The HI titers of groups inoculated with GM, YF18, and GZ289 were 8.0 log_2, 7.7 log_2, and 7.3 log_2, respectively. In the ducks in the contact group, the HI titers for YF18 were all 5 log_2, though none of the three ducks in the GM and GZ289 groups seroconverted (**Table 4**).

In the inoculated ducks, the YF18 virus was replicated systemically in the tissues of the thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen on 3 DPI (**Table 4**). The YF18 virus replicated more highly in the lungs (5.00 $log_{10}EID_{50}$). The mean virus titers in the thymus, cecal tonsils, bursa of Fabricius, trachea, brain, kidney, and spleen were 3.33, 2.67, 2.75, 1.92, 1.58, 3.58, and 2.83 $log_{10}EID_{50}$, respectively. Generally, the GM and GZ289 virus titers were lower than those of YF18 in ducks. Those two viruses replicated in some tested tissues, including those of the thymus, cecal tonsils, bursa of Fabricius, lung, kidney, and spleen, but not in those of the trachea or brain (**Table 4**). The GM and GZ289 viruses replicated more highly in the lungs (4.17 and 3.92 $log_{10}EID_{50}$, respectively).

GM, YF18, and GZ289 viruses shedding from the inoculated ducks were detected in oropharyngeal and cloacal swabs at 3, 5, 7, 9, and 11 DPI (**Table 5**). In the infected ducks, the GM virus could be detected only from oropharyngeal and cloacal swabs at 3 DPI (1.96 and 1.58 $log_{10}EID_{50}$, respectively). YF18 virus shedding was detected in the oropharynx in inoculated ducks within 7 DPI, with virus titers from 1.58 to 2.42 $log_{10}EID_{50}$, and from the cloaca within 3 DPI, with virus titers of 1.75 $log_{10}EID_{50}$. Lastly, GZ289 virus shedding occurred with both oropharyngeal and cloacal swabs at 3 DPI (1.75 and 1.83 $log_{10}EID_{50}$, respectively).

To determine whether these viruses could be horizontally transmitted among ducks, 8 h after infection three ducks were inoculated with 200 μL PBS via the same routes as a naïve contact group placed with those inoculated with the GM, YF18, or GZ289 virus. During the experiment period, no ducks in the naïve contact group inoculated with the GM, YF18, or GZ289 virus died (**Table 4**). In the naïve contact group representing YF18, the virus titers of oropharyngeal swabs were detectable only at 3 DPI (1.75 $log_{10}EID_{50}$), whereas cloacal swabs did not show any detectable virus during the period (**Table 5**). GM and GZ289 virus shedding was not testable in the oropharyngeal or cloacal swabs of the naïve contact duck, even at 14 DPI (**Table 5**).

In our study of the ducks, the YF18 virus was found to infect ducks and be transmitted among ducks via naïve contact.

TABLE 3 | Virus titers in oraopharyngeal and cloacal swabs from chickens.

Viruses		Days post-inoculation ($log_{10}EID_{50}$/0.1 mL) ± SD[a]									
		3 Days		5 Days		7 Days		9 Days		11 Days	
		OP	CL	OP	CL	OP	CL	OP	CL	OP	CL
GM	Infected[b]	4.79 ± 0.60 (6/6)	5.29 ± 0.40 (6/6)	–	–	–	-	–	–	–	–
	Contact[c]	4.13 ± 0.18 (2/3)	3.75 ± 0.43 (3/3)	4.00 (2/3)	3.92 ± 0.14(3/3)	–	–	–	–	–	–
YF18	Infected	5.25 ± 0.27(5/5)	4.92 ± 0.41(5/5)	2.50 (1/1)	ND (0/1)[d]	–	–	–	–	–	–
	Contact	4.00 ± 0.66 (3/3)	4.17 ± 0.14 (3/3)	3.25 ± 0.43(3/3)	3.25 ± 0.71(2/3)	3.00 (1/2)	3.00 (1/2)	–	–	–	–
GZ289	Infected	3.29 ± 0.62 (6/6)	2.88 ± 0.52 (6/6)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)

OP, oropharyngeal swabs; CL, cloacal swabs; –, all of the chickens died at the end of the observation.

[a]For statistical purposes, a value of 1.5 was assigned if virus was not detected from the undiluted sample in three embryonated hen's eggs (Kang et al., 2015).

[b]Chickens inoculated with virus.

[c]Naïve contact chickens housed with those inoculated.

[d]No detected.

Although the GM and GZ289 viruses could infect ducks, they could not be transmitted among them by naïve contact (**Figure 2**).

Pathogenicity and Transmission of NDV among Pigeons

To investigate the tissue tropism and pathogenicity of the three viruses in pigeons, we inoculated each pigeon with 10^6 EID_{50} of the indicated virus in 200 μL via intraocular and intranasal routes and euthanized three pigeons from each subgroup at 3 DPI. All remaining pigeons were observed clinically for 14 days.

No pigeons died during the observation period. In pigeons inoculated with the GZ289 virus, or in those in naïve contact with GZ289 virus-inoculated pigeons, HI titers were far higher than those observed for the two other viruses. In GZ289-inoculated pigeons, three seroconverted and showed high titers (9.3 log_2), whereas two pigeons in the contact group seroconverted with relatively high titers (6.0 log_2). The HI titers of groups inoculated with the GM and YF18 virus were 8.3 log2, and 8.7 log_2, respectively, although none of the three contact pigeons seroconverted (**Table 6**).

The GM and GZ289 viruses replicated systemically in pigeons, which was detectable in all tested tissues at 3 DPI, including those of the thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen (**Table 6**). The YF18 virus replicated only in some tested tissues, including those of the thymus, lung, cecal tonsils, kidney, and spleen; mean virus titers were 1.67, 2.33, 1.83, 1.75, and 2.08 $log_{10}EID_{50}$, respectively. The GM, YF18, and GZ289 viruses showed remarkable replication in the lungs, with mean titers of 4.42, 2.33, and 5.33 $log_{10}EID_{50}$, respectively. The three selected viruses furthermore replicated in the spleen to mean titers of 2.92, 2.08, and 4.92 $log_{10}EID_{50}$. These results indicate that GZ289 replicated more highly than the other two viruses in tested tissues of infected pigeons.

GM, YF18, and GZ289 viruses shedding from the inoculated pigeons were detected in oropharyngeal and cloacal swabs at 3, 5, 7, 9, and 11 DPI (**Table 7**). The GM virus could be isolated from both the oropharyngeal and cloacal swabs within 5 DPI (1.54–1.83 and 1.83–2.42 $log_{10}EID_{50}$, respectively). In the YF18 virus- inoculated group, the virus titers of the cloacal swabs were detectable only at 3 DPI (1.75 $log_{10}EID_{50}$); however, the virus titers of the oropharyngeal swabs could not be detected during our observation period (**Table 7**). The GZ289 virus was shed from both the oropharyngeal and cloacal swabs in inoculated pigeons within 11 DPI, except at 3 DPI (1.58–3.83 and 1.92–3.33 $log_{10}EID_{50,}$ respectively).

To determine whether these three viruses could be horizontally transmitted among pigeons, 8 h after infection, three pigeons were inoculated with 200 μL PBS via the same routes as a naïve contact group placed with those inoculated with the GM, YF18, or GZ289 viruses. No pigeons died in the naïve contact subgroup placed with those exposed to the three selected viruses (**Table 6**). No virus in the naïve contact pigeons with GM or YF18 could be isolated from oropharyngeal or cloacal swabs (**Table 7**). In the naïve contact group with GZ289, virus shedding could be detected from oropharyngeal swabs at 5, 7, and 9 DPI (1.67–1.83 $log_{10}EID_{50}$) and tested from cloacal swabs only at 7 and 9 DPI (1.58–1.67 $log_{10}EID_{50}$).

In sum, the GZ289 virus could infect and transmit among pigeons by naïve contact. Though the GM and YF18 viruses could infect pigeons, they could not transmit among pigeons by naïve contact (**Figure 2**).

DISCUSSION

South China's Guangdong Province is considered to be an ideal transmission area for NDV. It hosts numerous large-scale LBMs and live poultry markets, as well as a multitude of small backyard farms and small-scale poultry farms (Shortridge and Stuart-Harris, 1982). These poultry, including chickens, ducks, geese, pigeons, and numerous other species, are traded in LBMs daily. Due to the constant close proximity of these poultry, viruses achieve transmission in different birds and contribute to emergent novel NDVs. Indeed, South China is considered to be a virus epicenter, due to large-scale severe acute respiratory syndrome, high pathogenic avian influenza, H5N1, and dengue

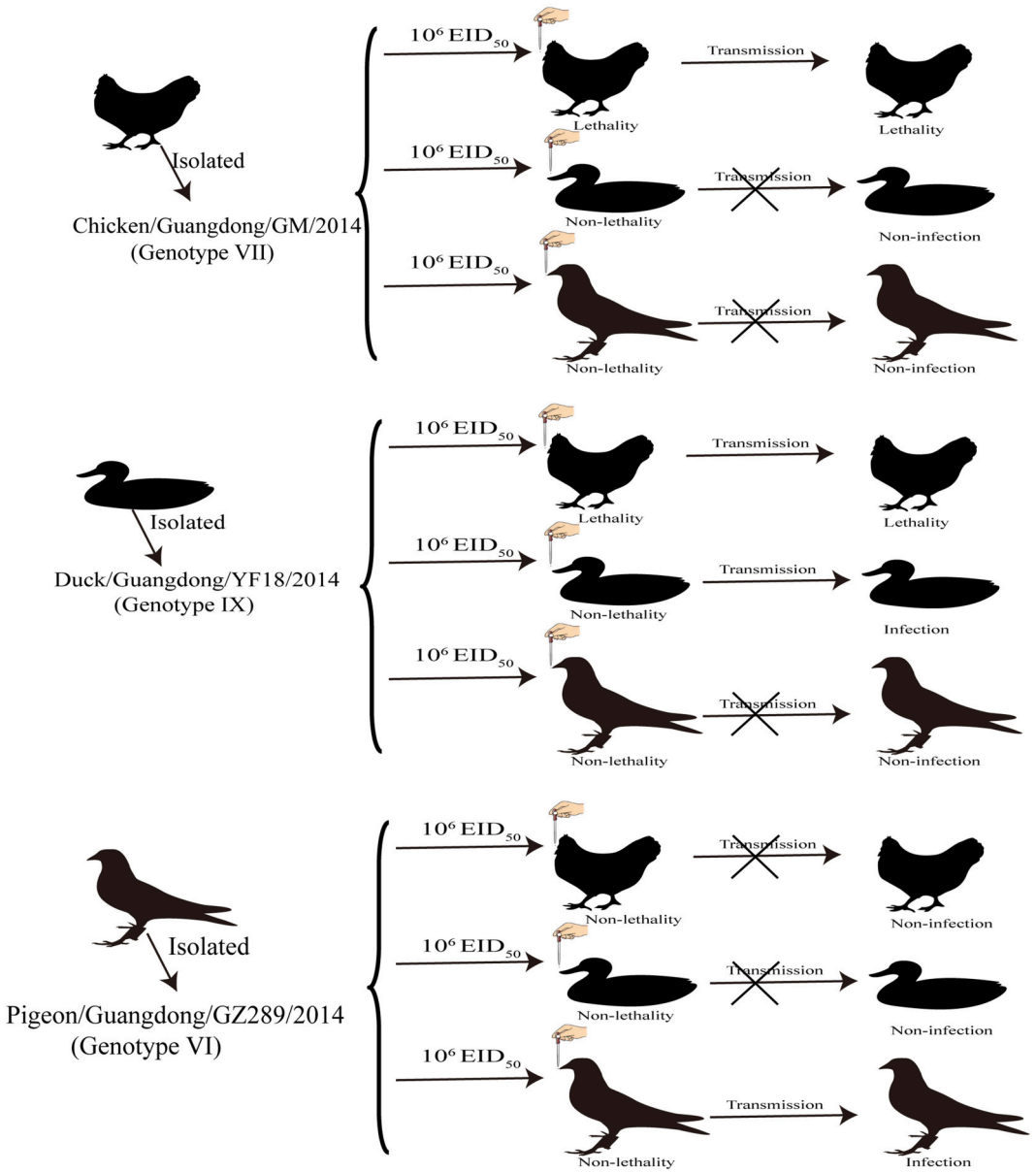


FIGURE 2 | Transmission studies of NDV in chickens, ducks and pigeons. Three groups of chickens, ducks, and pigeons were inoculated with 10^6 EID_{50} of the indicated virus via the intraocular and intranasal routes and, 8 h after infection, three birds were placed into the same cage to monitor contact infection.

outbreaks (Shortridge and Stuart-Harris, 1982; Qiu et al., 1993; Zhong et al., 2003; Chen et al., 2004). We therefore propose conducting routine surveillance, using chilled instead of live poultry for sale, and temporary rest days in poultry markets to prevent the intra- and interspecies transmission of NDV.

In recent years, NDV has caused large-scale outbreaks in poultry in many countries around the world, including China (Zhang et al., 2011; Chong et al., 2013; Kang et al., 2014), Japan (Mase et al., 2002), Southern Brazil (Marks et al., 2014), Indonesia (Xiao et al., 2012), South America (Diel et al., 2012), and West Malaysia (Jaganathan et al., 2015). In China, though the implementation of intensive vaccination and the culling of infected birds are effective policies for controlling ND in poultry and rural farms, virulent NDV can still frequently be isolated in vaccinated poultry (Liu et al., 2003; Zhang et al., 2011; Kang et al., 2014). Genetic and phylogenetic studies have shown that NDV is continuously evolving, with viruses of different genotypes undergoing simultaneous changes in distinct geographic areas (Diel et al., 2012; Chong et al., 2013). In our study, we characterized genetic and pathotypic properties of NDV strains isolated from chickens, ducks, and pigeons in LBMs in the province. The genetic and phylogenetic analysis of the complete sequences of the F protein gene showed that seven of 23 poultry-derived strains were avirulent class I NDV, three of 23 were class

TABLE 4 | Lethality, seroconversion, and tissues tropism among ducks in an intraspecies study of NDV transmission[a].

Viruses	Manifestations of ducks				Virus replication on 3 DPI ($log_{10}EID_{50}/g$)[d] in:							
	No. D/S/total[b]		No. S.C./total[c]									
	Infected[e]	Contact[f]	Infected	Contact	Thymus	Cecal tonsils	Bursa of Fabricius	Trachea	Lung	Brain	Kidney	Spleen
GM	0/0/3	0/0/3	3/3 (8.0)[g]	0/3	3.08 ± 0.29	2.58 ± 0.14	1.67 ± 0.14	ND	4.17 ± 0.14	ND	2.50 ± 0.25	2.42 ± 0.29
YF18	0/0/3	0/0/3	3/3 (7.7)	3/3 (5.0)	3.33 ± 0.14	2.67 ± 0.52	2.75 ± 0.43	1.92 ± 0.29	5.00 ± 0.43	ND	3.58 ± 0.14	2.83 ± 0.38
GZ289	0/0/3	0/0/3	3/3 (7.3)	0/3	2.75 ± 0.43	2.08 ± 0.38	1.58 ± 0.14	ND	3.92 ± 0.63	ND	2.83 ± 0.38	2.58 ± 0.58

[a] *Ducks were inoculated with 10^6 EID_{50} of the indicated virus in a volume of 200 μL via the intraocular and intranasal routes; Tissues, including thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen were collected aseptically on three DPI from three ducks when infected with the indicated virus and homogenized for virus titration in eggs.*

[b] *No. D/S/total shows the number of dead (D) and sick (S) as well as the total number of ducks from the observation period. The ducks that showed clinical signs, such as depression, torticollis, incoordination, and tremors, but recovered at the end of the observation were counted as sick animals.*

[c] *No. S.C./total shows the number of ducks that seroconverted out of the total number of ducks at the end of the observation period.*

[d] *For statistical analysis, a value of 1.5 was assigned if the virus was not detected from the undiluted sample in three SPF embryonated chicken eggs (Kang et al., 2015).Virus titers are expressed as means ± standard deviation in $log_{10}EID_{50}/g$ of tissue. ND, Not detected.*

[e] *Ducks inoculated with virus.*

[f] *Three uninoculated ducks were co-housed with infected ducks as a contact group 8 h after inoculation.*

[g] *Average antibody titer of infected ducks (log_2).*

TABLE 5 | Virus titers in oropharyngeal and cloacal swabs from ducks.

Viruses		Days post-inoculation ($log_{10}EID_{50}/0.1$ mL) ± SD[a]									
		3 Days		5 Days		7 Days		9 Days		11 Days	
		OP	CL	OP	CL	OP	CL	OP	CL	OP	CL
GM	Infected[b]	1.96 ± 0.71 (2/6)	1.58 ± 0.20 (1/6)	ND[d](0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact[c]	ND(0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
YF18	Infected	2.42 ± 1.01 (3/6)	1.75 ± 0.42 (2/6)	1.58 ± 0.14 (1/3)	ND (0/3)	1.58 ± 0.14(1/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact	1.75 ± 0.43 (1/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
GZ289	Infected	1.75 ± 0.39 (2/6)	1.83 ± 0.38 (3/6)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)

OP, oropharyngeal swabs; CL, cloacal swabs.

[a] *For statistical purposes, a value of 1.5 was assigned if virus was not detected from the undiluted sample in three embryonated hen's eggs (Kang et al., 2015).*

[b] *Ducks inoculated with virus.*

[c] *Naïve contact ducks housed with those inoculated.*

[d] *No detected.*

TABLE 6 | Lethality, seroconversion, and tissues tropism among pigeons in an intraspecies study of NDV transmission[a].

Viruses	Manifestations of pigeons				Virus replication on 3 DPI ($log_{10}EID_{50}/g$)[d] in:							
	No. D/S/total[b]		No. S.C./total[c]									
	Infected[e]	Contact[f]	Infected	Contact	Thymus	Cecal tonsils	Bursa of Fabricius	Trachea	Lung	Brain	Kidney	Spleen
GM	0/0/3	0/0/3	3/3 (8.3)[g]	0/3	3.08 ± 0.76	3.92 ± 0.52	2.33 ± 0.63	1.83 ± 0.58	4.42 ± 0.14	2.42 ± 0.29	3.08 ± 0.29	2.92 ± 0.29
YF18	0/0/3	0/0/3	3/3 (8.7)	0/3	1.67 ± 0.14	1.83 ± 0.38	ND	ND	2.33 ± 0.52	ND	1.75 ± 0.43	2.08 ± 0.52
GZ289	0/1/3	0/0/3	3/3 (9.3)	2/3 (6.0)	3.83 ± 0.58	4.33 ± 0.38	4.58 ± 0.14	2.83 ± 0.38	5.33 ± 0.38	2.75 ± 0.43	3.75	4.92 ± 1.01

[a] *Pigeons were inoculated with 10^6 EID_{50} of the indicated virus in a volume of 200 μL via the intraocular and intranasal routes; Tissues, including thymus, cecal tonsils, bursa of Fabricius, trachea, lung, brain, kidney, and spleen were collected aseptically on three DPI from three pigeons when infected with the indicated virus and homogenized for virus titration in eggs.*

[b] *No. D/S/total shows the number of dead (D) and sick (S) as well as the total number of pigeons from the observation period. The pigeons that showed clinical signs, such as depression, torticollis, incoordination, and tremors, but recovered at the end of the observation were counted as sick animals.*

[c] *No. S.C./total shows the number of pigeons that seroconverted out of the total number of pigeons at the end of the observation period.*

[d] *For statistical analysis, a value of 1.5 was assigned if the virus was not detected from the undiluted sample in three SPF embryonated chicken eggs (Kang et al., 2015).Virus titers are expressed as means ± standard deviation in $log_{10}EID_{50}/g$ of tissue. ND, Not detected.*

[e] *Pigeons inoculated with virus.*

[f] *Three uninoculated pigeons were co-housed with infected pigeons as a contact group 8 h after inoculation.*

[g] *Average antibody titer of infected pigeons (log_2).*

TABLE 7 | Virus titers in oropharyngeal and cloacal swabs from pigeons.

Viruses		Days post-inoculation ($log_{10}EID_{50}/0.1$ mL) ± SD[a]									
		3 Days		5 Days		7 Days		9 Days		11 Days	
		OP	CL	OP	CL	OP	CL	OP	CL	OP	CL
GM	Infected[b]	1.54 ± 0.10(1/6)	1.83 ± 0.52(2/6)	1.83 ± 0.38(2/3)	2.42 ± 1.01(2/3)	ND[d] (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact[c]	ND (0/3)	ND (0/3)	ND (0/3)	ND(0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
YF18	Infected	ND (0/3)	1.75 ± 0.42(2/6)	ND (0/3)	ND(0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
	Contact	ND (0/3)	ND (0/3)	ND (0/3)	ND(0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)	ND (0/3)
GZ289	Infected	ND (0/3)	ND (0/3)	2.00 ± 0.50(2/3)	2.83 ± 0.95(2/3)	3.83 ± 1.15(3/3)	3.33 ± 0.59(3/3)	2.79 ± 0.99(3/3)	2.71 ± 0.86(3/3)	1.58 ± 0.14(1/3)	1.92 ± 0.29(2/3)
	Contact	ND (0/3)	ND (0/3)	1.75 ± 0.43(2/3)	ND(0/3)	1.83 ± 0.58(2/3)	1.67 ± 0.14(2/3)	1.67 ± 0.29(1/3)	1.58 ± 0.14(1/3)	ND (0/3)	ND (0/3)

OP, oropharyngeal swabs; CL, cloacal swabs.

[a] *For statistical purposes, a value of 1.5 was assigned if virus was not detected from the undiluted sample in three embryonated hen's eggs (Kang et al., 2015).*

[b] *Pigeons inoculated with virus.*

[c] *Naïve contact pigeons housed with those inoculated.*

[d] *No detected.*

II genotype I, and one strain was class II genotype II. These results indicate that, similar to low pathogenic avian influenza, lentogenic NDVs prevalently circulate among domestic poultry at LBMs (Seal et al., 2005; Zhu et al., 2014). Additionally, as results of phylogenetic analyses show, 12 velogenic strains isolated from different birds related to predominant strains of class II genotypes VI, VII, and IX, thus suggesting the coexistence of different genotypes of NDV circulating simultaneously in South China, as well as a high probability of the emergence of new strains via recombination. Accordingly, epidemiological surveillance and further investigation at LBMs in South China is necessary in order to clarify the genetic evolution of NDV and thus issue early warnings.

PPMV-1 is generally virulent, though upon infecting chickens can result in clinical diseases expected of NDV with low virulence (OIE, 2012); however, these viruses remain a hidden threat to the poultry industry (de Oliveira Torres Carrasco et al., 2008). Previous studies have demonstrated that PPMV-1 strains are capable of being transmitted from infected pigeons to chickens and turkeys housed in physical contact, as well as that systemic replication can occur in those chickens and turkeys, as shown by the shedding the virus via oropharyngeal and cloacal routes and a humoral immune response to the virus; however quails and geese did not exhibit any clinical signs or shed the virus (Alexander and Parsons, 1984; Smietanka et al., 2014). Very few studies have examined the infectivity, pathogenesis, and transmission of NDV and PPMV-1 infections in different birds. In response, the aim of our study was to investigate the susceptibility and transmission of chickens, ducks, and pigeons following infection with three NDV strains—namely, Chicken/Guangdong/GM/2014 (GM), Duck/Guangdong/YF18/2014 (YF18), and Pigeon/Guangdong/GZ289/2014 (GZ289)—and to provide useful information for improving control strategies against ND. Our results demonstrate that GM is highly pathogenic to chickens and can transmit among them as well as ducks while circulating in chickens. YF18 was highly pathogenic in chickens, might have moderate or low pathogenicity in ducks and pigeons, and does not transmit to pigeons. In addition, GZ289 isolated from pigeons showed low pathogenicity to chickens and domestic ducks and could transmit only in pigeons (**Figure 2**). These results showed that NDVs isolated from different birds exhibit different host ranges and tissue tropisms. Nevertheless, our study posed several limitations—for instance, we do not know the infective dose for each virus for each species, owing to the adaptability of a virus within a single species.

At least one previous study has reported that pigeons exhibited high morbidity and mortality rates, whereas chickens showed no clinical signs, when infected with the same PPMV-1 strain (Guo et al., 2014). However, opposite results were found by Dortmans et al. (2011)—ones consistent with the results of our experiments—who failed to induce clinical signs in pigeons infected with pigeon strain AV324 or FL-Herts, though the virus was shed from the oropharynx and cloaca in inoculated pigeons (Dortmans et al., 2011). These findings suggest that the course of experimental infection with PPMV-1 in different birds can vary greatly and most likely depends on the infective dose for each of the viruses, the inoculation route, the immunity of the host, and the age and species of the birds.

Current NDV vaccines in circulation, including class II genotype II vaccine virus (B1, Clone-30, and La Sota) and genotype I vaccine virus (V4), are still used at a large scale, most extensively for protecting poultry flocks from ND in South China (Hu et al., 2009). However, until now, well-controlled findings have not demonstrated the role of vaccination in any attempt to control NDV outbreaks by preventing virus transmission in poultry flocks. Moreover, current vaccines can prevent NDV outbreaks, yet not stop viral shedding in vaccinated poultry flocks (Kapczynski and King, 2005). Additional studies are therefore needed to identify the best vaccine candidate, not only for preventing clinical disease and mortality, but also to decrease the magnitude of viral shedding from vaccinated birds.

A correlation exists between antibody response and shedding after infection with virulent NDV in susceptible animals (Miller et al., 2013). During the course of our study, the statistical analysis of serological results showed significant differences among chicken, duck, and pigeon groups exposed to different viruses, as well as the naïve contacts. In chickens in the contact group, the virus was detectable from oropharyngeal and cloacal swabs inoculated with GM and YF18, whereas in ducks in the contact group, we could detect the virus only from oropharyngeal and cloacal swabs inoculated with YF18, with HI titers for the GM, YF18, and GZ289 of 5, 6,and 4 $\log_2$, respectively. In contact group pigeons, the virus could be detected only from oropharyngeal and cloacal swabs inoculated with GZ289. Moreover, HI titers for GZ289 at 14 DPI were all 6 $\log_2$, though none of the three pigeons in the GM and YF18 groups seroconverted (HI titers = 4). In all, the efficient replication, high seroconversion, and shedding of relatively high titers in naïve contact groups suggest that NDV isolated from different birds was transmitted to the naïve contact group.

Altogether, our results provide clear evidence that genetically diverse viruses circulate in LBMs in South China's Guangdong Province and illustrate that the three NDV strains isolated from different birds have varying levels of infectivity, pathogenicity, and transmission in chickens, ducks, and pigeons. Our findings thus emphasize the need for constant epidemiological studies in LBMs, in order to enhance active surveillance toward preventing the spread and evolution of these viruses.

AUTHOR CONTRIBUTIONS

YK conceived the study and wrote the paper. BX and YK designed, performed, and analyzed all the experiments. RY provided technical assistance and prepared all the figures. YL, SF, YL, and TR designed the study and revised the manuscript. All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This study was mainly funded by grants from the National Natural Science Foundation of China (No.

31372412), the Chinese Special Fund for Agro-scientific Research in the Public Interest (No.201303033), the Science and Technology Projects of Guangdong Province (No.2012A020800006), and the Specialized Research Fund for Doctoral Program of Higher Education of China (No. 20124404 110016).

REFERENCES

Alexander, D. J. (2000). Newcastle disease and other avian paramyxoviruses. *Rev. Sci. Tech.* 19, 443–462. Available online at: http://uganda.totoagriculture.org/PDFs/LivestockDiseases/485.pdf

Alexander, D. J., and Parsons, G. (1984). Avian paramyxovirus type 1 infections of racing pigeons: 2 pathogenicity experiments in pigeons and chickens. *Vet. Rec.* 114, 466–469. doi: 10.1136/vr.114.19.466

Ballagi-Pordány, A., Wehmann, E., Herczeg, J., Belák, S., and Lomniczi, B. (1996). Identification and grouping of Newcastle disease virus strains by restriction site analysis of a region from the F gene. *Arch. Virol.* 141, 243–261. doi: 10.1007/BF01718397

Bogoyavlenskiy, A., Berezin, V., Prilipov, A., Usachev, E., Lyapina, O., Levandovskaya, S., et al. (2005). Molecular characterization of virulent Newcastle disease virus isolates from chickens during the 1998 NDV outbreak in Kazakhstan. *Virus Genes* 31, 13–20. doi: 10.1007/s11262-004-2195-2

Cai, S., Li, J., Wong, M. T., Jiao, P., Fan, H., Liu, D., et al. (2011). Genetic characterization and evolutionary analysis of 4 Newcastle disease virus isolate full genomes from waterbirds in South China during 2003-2007. *Vet. Microbiol.* 152, 46–54. doi: 10.1016/j.vetmic.2011.04.014

Cattoli, G., Susta, L., Terregino, C., and Brown, C. (2011). Newcastle disease: a review of field recognition and current methods of laboratory detection. *J. Vet. Diagn. Invest.* 23, 637–656. doi: 10.1177/1040638711407887

Chen, H., Deng, G., Li, Z., Tian, G., Li, Y., Jiao, P., et al. (2004). The evolution of H5N1 influenza viruses in ducks in southern China. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10452–10457. doi: 10.1073/pnas.0403212101

Chong, Y. L., Lam, T. T., Kim, O., Lu, H., Dunn, P., and Poss, M. (2013). Successful establishment and global dispersal of genotype VI avian paramyxovirus serotype 1 after cross species transmission. *Infect. Genet. Evol.* 17, 260–268. doi: 10.1016/j.meegid.2013.04.025

de Oliveira Torres Carrasco, A., Seki, M. C., de Freitas Raso, T., Paulillo, A. C., and Pinto, A. A. (2008). Experimental infection of Newcastle disease virus in pigeons (*Columba livia*): humoral antibody response, contact transmission and viral genome shedding. *Vet. Microbiol.* 129, 89–96. doi: 10.1016/j.vetmic.2007.11.012

Diel, D. G., Susta, L., Cardenas Garcia, S., Killian, M. L., Brown, C. C., Miller, P. J., et al. (2012). Complete genome and clinicopathological characterization of a virulent Newcastle disease virus isolate from South America. *J. Clin. Microbiol.* 50, 378–387. doi: 10.1128/JCM.06018-11

Dortmans, J. C., Koch, G., Rottier, P. J., and Peeters, B. P. (2011). A comparative infection study of pigeon and avian paramyxovirus type 1 viruses in pigeons: evaluation of clinical signs, virus shedding and seroconversion. *Avian Pathol.* 40, 125–130. doi: 10.1080/03079457.2010.542131

Erickson, G., Brugh, M., and Beard, C. (1980). Viscerotropic velogenic Newcastle disease in pigeons: clinical disease and immunization. *Avian Dis.* 24, 257–267. doi: 10.2307/1589785

Guo, H., Liu, X., Xu, Y., Han, Z., Shao, Y., Kong, X., et al. (2014). A comparative study of pigeons and chickens experimentally infected with PPMV-1 to determine antigenic relationships between PPMV-1 and NDV strains. *Vet. Microbiol.* 168, 88–97. doi: 10.1016/j.vetmic.2013.11.002

Hu, S., Ma, H., Wu, Y., Liu, W., Wang, X., Liu, Y., et al. (2009). A vaccine candidate of attenuated genotype VII Newcastle disease virus generated by reverse genetics. *Vaccine* 27, 904–910. doi: 10.1016/j.vaccine.2008.11.091

Jaganathan, S., Ooi, P. T., Phang, L. Y., Allaudin, Z. N., Yip, L. S., Choo, P. Y., et al. (2015). Observation of risk factors, clinical manifestations and genetic characterization of recent Newcastle Disease Virus outbreak in West Malaysia. *BMC Vet. Res.* 11:219. doi: 10.1186/s12917-015-0537-z

Jinding, C., Ming, L., Tao, R., and Chaoan, X. (2005). A goose-sourced paramyxovirus isolated from southern China. *Avian Dis.* 49, 170–173. doi: 10.1637/7270-090101R1

Kaleta, E. F., and Baldauf, C. (1988). "Newcastle disease in free-living and pet birds," in *Newcastle Disease*, ed D. J. Alexander (Boston, MA: Springer), 197–246.

Kang, Y., Li, Y., Yuan, R., Feng, M., Xiang, B., Sun, M., et al. (2015). Host innate immune responses of ducks infected with Newcastle Disease Viruses of different pathogenicities. *Front. Microbiol.* 6:1283. doi: 10.3389/fmicb.2015.01283

Kang, Y., Li, Y., Yuan, R., Li, X., Sun, M., Wang, Z., et al. (2014). Phylogenetic relationships and pathogenicity variation of two Newcastle disease viruses isolated from domestic ducks in Southern China. *Virol. J.* 11:147. doi: 10.1186/1743-422X-11-147

Kapczynski, D. R., and King, D. J. (2005). Protection of chickens against overt clinical disease and determination of viral shedding following vaccination with commercially available Newcastle disease virus vaccines upon challenge with highly virulent virus from the California 2002 exotic Newcastle disease outbreak. *Vaccine* 23, 3424–3433. doi: 10.1016/j.vaccine.2005.01.140

Kim, L. M., King, D. J., Curry, P. E., Suarez, D. L., Swayne, D. E., Stallknecht, D. E., et al. (2007). Phylogenetic diversity among low-virulence newcastle disease viruses from waterfowl and shorebirds and comparison of genotype distributions to those of poultry-origin isolates. *J. Virol.* 81, 12641–12653. doi: 10.1128/JVI.00843-07

Kommers, G. D., King, D. J., Seal, B. S., and Brown, C. C. (2003). Virulence of six heterogeneous-origin Newcastle disease virus isolates before and after sequential passages in domestic chickens. *Avian Pathol.* 32, 81–93. doi: 10.1080/0307945021000070750

Liu, X. F., Wan, H. Q., Ni, X. X., Wu, Y. T., and Liu, W. B. (2003). Pathotypical and genotypical characterization of strains of Newcastle disease virus isolated from outbreaks in chicken and goose flocks in some regions of China during 1985-2001. *Arch. Virol.* 148, 1387–1403. doi: 10.1007/s00705-003-0014-z

Marks, F. S., Rodenbusch, C. R., Okino, C. H., Hein, H. E., Costa, E. F., Machado, G., et al. (2014). Targeted survey of Newcastle disease virus in backyard poultry flocks located in wintering site for migratory birds from Southern Brazil. *Prev. Vet. Med.* 116, 197–202. doi: 10.1016/j.prevetmed.2014.06.001

Mase, M., Imai, K., Sanada, Y., Sanada, N., Yuasa, N., Imada, T., et al. (2002). Phylogenetic analysis of Newcastle disease virus genotypes isolated in Japan. *J. Clin. Microbiol.* 40, 3826–3830. doi: 10.1128/JCM.40.10.3826-3830.2002

Miller, P. J., Afonso, C. L., El Attrache, J., Dorsey, K. M., Courtney, S. C., Guo, Z., et al. (2013). Effects of Newcastle disease virus vaccine antibodies on the shedding and transmission of challenge viruses. *Dev. Comp. Immunol.* 41, 505–513. doi: 10.1016/j.dci.2013.06.007

Miller, P. J., Decanini, E. L., and Afonso, C. L. (2010). Newcastle disease: evolution of genotypes and the related diagnostic challenges. *Infect. Genet. Evol.* 10, 26–35. doi: 10.1016/j.meegid.2009.09.012

OIE. (2012). "Newcastle disease, chapter 2.3.14," in *OIE Terrestrial Manual 2012: Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (Paris: World Organisation for Animal Health), 576–589. Available online at: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.03.14_NEWCASTLE_DIS.pdf

Qiu, F. X., Gubler, D. J., Liu, J. C., and Chen, Q. Q. (1993). Dengue in China: a clinical review. *Bull. World Health Organ.* 71, 349–359.

Qiu, X., Sun, Q., Wu, S., Dong, L., Hu, S., Meng, C., et al. (2011). Entire genome sequence analysis of genotype IX Newcastle disease viruses reveals their early-genotype phylogenetic position and recent-genotype genome size. *Virol. J.* 8:117. doi: 10.1186/1743-422X-8-117

Reed, L. J., and Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* 27, 493–497.

Seal, B. S., Wise, M. G., Pedersen, J. C., Senne, D. A., Alvarez, R., Scott, M. S., et al. (2005). Genomic sequences of low-virulence avian paramyxovirus-1 (Newcastle disease virus) isolates obtained from live-bird markets in North

America not related to commonly utilized commercial vaccine strains. *Vet. Microbiol.* 106, 7–16. doi: 10.1016/j.vetmic.2004.11.013

Shengqing, Y., Kishida, N., Ito, H., Kida, H., Otsuki, K., Kawaoka, Y., et al. (2002). Generation of velogenic Newcastle disease viruses from a nonpathogenic waterfowl isolate by passaging in chickens. *Virology* 301, 206–211. doi: 10.1006/viro.2002.1539

Shortridge, K. F., and Stuart-Harris, C. H. (1982). An influenza epicentre? *Lancet* 2, 812–813.

Sinkovics, J. G., and Horvath, J. C. (2000). Newcastle disease virus (NDV): brief history of its oncolytic strains. *J. Clin. Virol.* 16, 1–15. doi: 10.1016/S1386-6532(99)00072-4

Smietanka, K., Olszewska, M., Domanska-Blicharz, K., Bocian, A. L., and Minta, Z. (2014). Experimental infection of different species of birds with pigeon paramyxovirus type 1 virus–evaluation of clinical outcomes, viral shedding, and distribution in tissues. *Avian Dis.* 58, 523–530. doi: 10.1637/10769-011514-Reg.1

Snoeck, C. J., Owoade, A. A., Couacy-Hymann, E., Alkali, B. R., Okwen, M. P., Adeyanju, A. T., et al. (2013). High genetic diversity of Newcastle disease virus in poultry in West and Central Africa: cocirculation of genotype XIV and newly defined genotypes XVII and XVIII. *J. Clin. Microbiol.* 51, 2250–2260. doi: 10.1128/JCM.00684-13

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Wang, J., Liu, H., Liu, W., Zheng, D., Zhao, Y., Li, Y., et al. (2015). Genomic characterizations of six pigeon paramyxovirus type 1 viruses isolated from live bird markets in China during 2011 to 2013. *PLoS ONE* 10:e0124261. doi: 10.1371/journal.pone.0124261

Wang, Z., Ning, Z., Sun, M., Gao, S., Kang, Y., Xie, P., et al. (2014). Interferon regulatory factor 7- (IRF7-) mediated immune response affects Newcastle disease virus replication in chicken embryo fibroblasts. *Acta Vet. Hung.* 62, 500–511. doi: 10.1556/AVet.2014.023

Xiao, S., Paldurai, A., Nayak, B., Samuel, A., Bharoto, E. E., Prajitno, T. Y., et al. (2012). Complete genome sequences of Newcastle disease virus strains circulating in chicken populations of Indonesia. *J. Virol.* 86, 5969–5970. doi: 10.1128/JVI.00546-12

Zhang, S., Wang, X., Zhao, C., Liu, D., Hu, Y., Zhao, J., et al. (2011). Phylogenetic and pathotypical analysis of two virulent Newcastle disease viruses isolated from domestic ducks in China. *PLoS ONE* 6:e25000. doi: 10.1371/journal.pone.0025000

Zhong, N. S., Zheng, B. J., Li, Y. M., Poon, Xie, Z. H., Chan, K. H., et al. (2003). Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* 362, 1353–1358. doi: 10.1016/S0140-6736(03)14630-2

Zhu, J., Xu, H., Liu, J., Zhao, Z., Hu, S., Wang, X., et al. (2014). Surveillance of avirulent Newcastle disease viruses at live bird markets in Eastern China during 2008-2012 reveals a new sub-genotype of class I virus. *Virol. J.* 11:211. doi: 10.1186/s12985-014-0211-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Kang, Xiang, Yuan, Zhao, Feng, Gao, Li, Li, Ning and Ren. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 October 2016
doi: 10.3389/fmicb.2016.01587

Infection of Goose with Genotype VIId Newcastle Disease Virus of Goose Origin Elicits Strong Immune Responses at Early Stage

Qianqian Xu[1,2], Yuqiu Chen[1,2], Wenjun Zhao[1,2], Tingting Zhang[1,2], Chenggang Liu[1,2], Tianming Qi[1,2], Zongxi Han[2], Yuhao Shao[2], Deying Ma[1] and Shengwang Liu[2]**

[1] College of Animal Science and Technology, Northeast Agricultural University, Harbin, China, [2] Division of Avian Infectious Diseases, State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin, China

Newcastle disease (ND), caused by virulent strains of Newcastle disease virus (NDV), is a highly contagious disease of birds that is responsible for heavy economic losses for the poultry industry worldwide. However, little is known about host-virus interactions in waterfowl, goose. In this study, we aim to characterize the host immune response in goose, based on the previous reports on the host response to NDV in chickens. Here, we evaluated viral replication and mRNA expression of 27 immune-related genes in 10 tissues of geese challenged with a genotype VIId NDV strain of goose origin (go/CH/LHLJ/1/06). The virus showed early replication, especially in digestive and immune tissues. The expression profiles showed up-regulation of Toll-like receptor (TLR)1–3, 5, 7, and 15, avian β-defensin (AvBD) 5–7, 10, 12, and 16, cytokines [interleukin (IL)-8, IL-18, IL-1β, and interferon-γ], inducible NO synthase (iNOS), and MHC class I in some tissues of geese in response to NDV. In contrast, NDV infection suppressed expression of AvBD1 in cecal tonsil of geese. Moreover, we observed a highly positive correlation between viral replication and host mRNA expressions of TLR1-5 and 7, AvBD4-6, 10, and 12, all the cytokines measured, MHC class I, FAS ligand, and iNOS, mainly at 72 h post-infection. Taken together, these results demonstrated that NDV infection induces strong innate immune responses and intense inflammatory responses at early stage in goose which may associate with the viral pathogenesis.

Keywords: NDV, goose, AvBD, TLR, cytokines, iNOS

OPEN ACCESS

Edited by:
Akio Adachi,
University of Tokushima, Japan

Reviewed by:
Takamasa Ueno,
Kumamoto University, Japan
Xiufan Liu,
Yangzhou University, China

****Correspondence:***
Deying Ma
madeying@neau.edu.cn
Shengwang Liu
swliu@hvri.ac.cn

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *19 July 2016*
Accepted: *22 September 2016*
Published: *04 October 2016*

Citation:
Xu Q, Chen Y, Zhao W, Zhang T, Liu C, Qi T, Han Z, Shao Y, Ma D and Liu S (2016) Infection of Goose with Genotype VIId Newcastle Disease Virus of Goose Origin Elicits Strong Immune Responses at Early Stage. Front. Microbiol. 7:1587. doi: 10.3389/fmicb.2016.01587

INTRODUCTION

Newcastle disease (ND), caused by Newcastle disease virus (NDV), is regarded as one of the most important avian diseases (Häuslaigner et al., 2009). The virus caused an economically serious disease in almost all poultry (Alexander, 1988; Sun et al., 2013). Waterfowls, such as duck and goose, are generally considered to be natural reservoirs or carriers of NDV, even those most virulent for chicken (Alexander, 2001; Alexander and Senne, 2008). However, ND outbreaks in domestic waterfowl have frequently been reported in East Asian countries, including Korea, Japan, and China, since the 1980s (Liu et al., 2003; Lee et al., 2004; Mase et al., 2009). In the affected flocks of duck, egg production sharply declined by about 70%, morbidity was about 80%, and mortality varied from

30 to 50%. The diseased birds showed diarrhea and nervous signs, and the dead birds mainly manifested by focal hemorrhage, necrosis of the intestinal mucosa, and congestion and hemorrhage of the ovarian follicles (Liu et al., 2010, 2015; Dai et al., 2013). Similarly, serious ND outbreaks have also been reported in flocks of geese in China (Liu et al., 2003). The geese challenged with NDV of goose origin showed clinical signs such as anorexia, white diarrhea, depression, nasal discharges, ocular, and dead (Wan et al., 2004; Häuslaigner et al., 2009). In Asia, the main virulent NDV strain belongs to Genotype VIId, which is a major threat to the poultry industry, including goose (Hu et al., 2015). Thus, understanding the underlying mechanism of NDV pathogenesis, as well as development of novel alternative therapeutic approaches is of great importance. Recently, most reports have focused on the molecular characteristics and pathogenicity of epidemic strains (Kumar et al., 2011; Chen et al., 2012; Wang et al., 2015), while the host–virus interactions remain largely unknown.

Innate immunity is considered to be the first line of host defense against virus infection. The host innate response play an important role early in the process on an infection, as some of these responses may prevent the initial viral replication or they may send appropriate signals in order to initiate other innate mechanisms as well as adaptive responses. Prominent among these are Toll-like receptors (TLRs) (Abdul-Careem et al., 2009). In mammals, at least 13 TLRs have been identified. In chicken, TLR1A and B, TLR2A and B, TLR3, TLR4, TLR5, TLR7, TLR15, and TLR21 have been identified (Paul et al., 2013). It has been shown that several TLRs recognize viral pathogen-associated molecular patterns (PAMPs): TLR3, detects double-stranded RNA (dsRNA) derived from viral replication whereas single stranded RNA (ssRNA) are detected by TLR7 and TLR8 (Sang et al., 2008). TLRs provide various means of limiting virus replication until adaptive immune responses are activated. For example, TLR3 plays a critical role in eliminating herpes simplex virus-2 infection in mouse reproductive tract (Ashkar et al., 2004; Abdul-Careem et al., 2009), or inhibiting NDV replication in HeLa cells (Cheng et al., 2014). Upon recognize invading virus, TLRs activate downstream signaling cascades, leading to the secretion of soluble factors, such as cytokines and host defense peptides (HDPs), which in turn mediate innate immune responses to limit viral replication (Seth et al., 2006; Takeuchi and Akira, 2010).

Newcastle disease virus induced host immune responses in chickens (Marina and Hanson, 1987). NO production *in vitro* was upregulated in chicken peripheral blood mononuclear cells (PBMCs) and heterophils in response to NDV challenge (Sick et al., 2000; Ahmed et al., 2007). Accordingly, inducible NO synthase (iNOS) mRNA expression was upregulated in NDV-infected chicken PBMCs *in vitro* (Ahmed et al., 2007). In addition, NDV also induced interferon (IFN)-α and IFN-β mRNA in chicken macrophages (Sick et al., 1998), IFN-γ mRNA in PBMCs (Ahmed et al., 2007), and IFN-α, IFN-β, interleukin (IL)-1β and IL-6 in chicken splenic leukocytes (Rue et al., 2011). In agreement with the observations *in vitro*, there are multiple genes induced in spleen of NDV-infected chickens *in vivo*, including some chemokines and cytokines, types I and II IFNs, IFN effectors and iNOS (Rue et al., 2011). These findings were confirmed by our recent results that pigeon NDV induces immune responses characterized by activation of TLRs, particularly TLR3 and TLR7, avian β-defensin (AvBD) 2 and 10, and iNOS of pigeons at 7 days post infection (Li et al., 2015). Up to now, particular attentions are mostly given to the pathogenicity and origin of NDV, while studies on host immune response to NDV infection are scarce. Here, a genotype VIId of class II NDV strain of goose origin (go/CH/LHLJ/1/06) (Xu et al., accepted), which was isolated from geese in 2006 in our laboratory, was selected as the model pathogen. In order to characterize the host immune response in goose, based on the previous reports on the host response to NDV in chickens, we examined the virus replication and induction of innate host responses in tissues of geese challenged with the NDV strain in this study.

MATERIALS AND METHODS

Ethics Statement

All animal experimental procedures were approved by the Ethical and Animal Welfare Committee of Heilongjiang Province, China (License no. SQ20150508).

Virus and Animals

Goose isolate of NDV (go/CH/LHLJ/1/06) was isolated from the field in 2006 in Heilongjiang Province, China, and proved to be a virulent NDV strain with mean dead time of 51 h in embryonating chicken eggs, and intracerebral pathogenicity index of 1.86 (Alexander et al., 1998; Xu et al., accepted). In addition, it was found that the virulent NDV strain causes a mortality rate of 20% in geese (Xu et al., accepted).

One-day-old geese were hatched in this study. The geese were reared to observe health status for 40 days old. Clinical health of the geese was confirmed by a histopathological examination. In addition, the sera of the birds were confirmed to be negative for NDV-specific hemagglutination inhibition antibodies before experiments were done. The challenge test was conducted in isolators in biosafety level 3 facilities under negative pressure.

Nucleotide Sequencing of Immune Molecules in Geese

The total RNA was extracted and cDNA was synthesized as previously described (Li et al., 2015). The resultant cDNA product was used for subsequent PCR by using Ex-Taq polymerase (Takara Bio, Shiga, Japan). The primers were designed by aligning the nucleotide sequences of respective genes from geese (*Anser*), chickens (*Gallus gallus*), and ducks (*Anas platyrhynchos*) using software of DNAStar (Lasergene Corp, Madison, WI, USA). Briefly, the primers of 18S rRNA, TLRs (2, 3, 4, 5, 7, and 15), AvBDs (1, 2, 3, 5, 6, 9, and 10), MHC class I, and cytokines (IL-1β, IL-2, IL-6, IL-8, IL-18, and IFN-γ), were designed based on respective nucleotide sequences of geese. The nucleotide sequences of TLR1, AvBDs (4, 7, 12, and 16), iNOS, and FAS ligand (FASLG) of both chickens (*G. gallus*) and ducks (*A. platyrhynchos*) were aligned by using software of DNAstar as

well. The consensus sequences were selected and used for primer design (**Table 1**). The PCR products were cloned and sequenced as previously described (Li et al., 2015). The resultant plasmids were used as respect control for subsequent quantitative RT-PCR.

Bioinformatic Analysis

Basic searches were conducted with BLAST program analysis[1]. Phylogenetic tree was performed from aligned amino acid sequences by the neighbor-joining method with 1,000 bootstraps using the MEGA4 program software (Molecular Evolutionary Genetics Analysis, version 6.0, Armonk, NY, USA).

Experimental Design and Real-Time RT-PCR of mRNAs in Tissues

At the age of 40 days, the geese were allotted randomly to three groups. Groups 1 and 2 have 10 birds each and was inoculated intranasally with 100 μl of the NDV strain (go/CH/LHLJ/1/06) at 10^6 50% egg infective doses. Group 3 has twenty birds and was inoculated with 100 μl of phosphate-buffered saline only, and served as a control. At 36 and 72 h post infection (hpi), five geese from Groups 1 and 3 were killed and ten tissue samples, including brain, trachea, lungs, kidneys, liver, proventriculus, spleen, cecal tonsil, Harderian glands, and bursa of Fabricius were collected. All of these collected tissues were used for real-time RT-PCR analysis of NDV and host genes. Birds in group 2 and the remaining birds in group 3 were kept to observe the clinical signs and mortality for 3 weeks post-challenge. The dead birds were examined for the gross lesions in different organs.

One-step Real-time PrimeScript RT-PCR (Takara Biotechnology, Dalian, China) was used to evaluate the mRNA levels of the selected genes as described previously (Li et al., 2015). The preparation of the real-time RT-PCR followed the QIME requirement http://www.clinchem.org/content/55/4/611.long. Briefly, RNA was extracted as described above. The assays were performed using 2 μL of total RNA and the One-step Real-time PrimeScript® RT-PCR kit (Takara Biotechnology, Dalian Co., Ltd.) in a 20-μL reaction on a LightCycler® 480II Real-Time PCR system (Roche, Basel, Switzerland) according to previous studies (Li et al., 2015; Xu et al., 2015). Serial tenfold dilution of plasmids containing goose18S rRNA, TLRs (1, 2, 3, 5, 7, and 15), AvBDs (1–7, 9, 10, 12, and 16), cytokines (IL-1β, IL-6, IL-8, and IL-18), iNOS, MHC class I and FASLG were used as controls. The mRNA of these genes was evaluated. All amplifications were conducted in triplicate. The concentration of target cDNA in a sample was deduced from the crossing point obtained and from the corresponding standard curve. The data are expressed for each sample as the copy number of each target cDNA normalized to that of the reference gene (18S rRNA), as described previously (Li et al., 2015).

One-step Real-time PrimeScript RT-PCR was also used for detecting the viral RNA of the NDV strain in the tissues of geese as described previously (Guo et al., 2014). The primers and probe were designed based on the *M* gene sequence of the NDV strain used in this study. The primers and probe were as follows:

forward, 5′–CTCAGTGATGTGCTCGGACC–3′;
reverse, 5′–CCTGGGGAGAGGCATTTGCTA–3′;
probe, 5′–[FAM]TTCTCTAGCAGTGGGACAGCCTGC[BHQ1]–3′ BH.

Statistical Analysis

Data are expressed as means ± SD. The statistical significance was assessed by using SAS software (1996) as previously described (Li et al., 2015). Correlation between the relative gene expression of immune molecules and NDV was performed using Pearson's tau using SAS software (1996), and $P < 0.05$ was considered to be statistically significant.

The nucleotide sequences of both of anser_AvBD7 and anser_AvBD12 obtained in current study are available from GenBank under the accession numbers KR018386 (anser_AvBD7) and KR018387 (anser_AvBD12). The nucleotide sequences of anser_AvBD4, anser_AvBD16, TLR1, FASLG and iNOS are shown in detail in Supplementary Figure S2–S6 in the Supplementary Materials, due to nucleotide sequences of them are shorter than 200 bp.

RESULTS

Sequence Analysis of Immune Molecules

Using a set of primers designed to amplify conserved AvBDs sequences, four novel AvBDs (4, 7, 12, and 16) were identified from both spleen and bone marrow of healthy geese. The open reading frames (ORFs) of two of these novel AvBDs contained 201 and 198 nt, respectively, and encoded 66 and 65 amino acids, respectively. The other two contained only parts of the ORFs with 171 and 180 bp, and encoded 56 and 59 amino acids, respectively. A BLASTN search revealed that the sequence of the first peptide (56 amino acids) showed the highest amino acid identity (80.9%) to chicken AvBD4, comparing with other AvBDs and defensins of the mammals. Hence, the first peptide was designed as goose AvBD4. The sequence of the remaining peptides (66, 65, and 59 amino acids) shared the highest amino acid identities to chicken AvBD7 (84.4%), AvBD12 (87.4%) and duck AvBD16 (66.3%), respectively, and designed as goose AvBD7, AvBD12, and AvBD16, respectively. Moreover, the GXC motif and the six cysteine residues were found in the predicted amino acid sequences of these four peptides that are conserved across all β-defensins. These four novel AvBDs were named anser_AvBD4, anser_AvBD7, anser_AvBD12, and anser_AvBD16 (**Figure 1**). Phylogenetic analyses were performed on the amino acid sequences of β-defensins, including these four novel AvBDs, the other reported AvBDs, and some mammalian β-defensins (Supplementary Figure S1). All these β-defensins segregated into eight distinct clades. Anser_AvBD4 formed a branch with AvBD4 from other avian species; anser_AvBD7 formed a branch with AvBD7 and AvBD6 from other avian species; anser_AvBD12 formed a branch with AvBD12–14 from chickens and two mammalian β-defensins (*Mus musculus* Defb2 and *Sus scrofa* PBD2); and anser_AvBD16 formed a branch with AvBD3, 8, 16, 103a, and 103b from other avian species.

[1]http://www.ncbi.nlm.nih.gov/blast

TABLE 1 | PCR primer sequences and predicted product lengths.

Target mRNA	Sense primer (5′–3′)	Antisense primer (5′–3′)	Product size (bp)	GenBank accession no.
AvBD4 (RT-PCR)	5′-ATCGTGCTCCTCTTTGTGGCAGTTCA-3′	5′-CTACAACCATCTACAGCAAGAATACT-3′	171	–
AvBD7 (RT-PCR)	5′-GGAGACAGAAGGCAGCGGTGAT -3′	5′-GGAGACAGAAGGCAGCGGTGAT -3′	326	KR018386
AvBD12 (RT-PCR)	5′-CAGCCCTGCTGCTCYCCAGCAG -3′	5′-TCAGGTCTTGGTGGGAGTTGGTGACAGAGG TTTACAGCAGAGAATGAC -3′	230	KR018387
AvBD16 (RT-PCR)	5′-ATGAAGATCCTCTGCCTGCTCTTC -3′	5′-TTAGTCGTACACAGTCCG(T)GCAGCACGGGAC -3′	180	–
AvBD1 (real time RT-PCR)	5′-GAAACAAGGAGAAATGTCATCG -3′	5′-ATGGGGGTTGTTTCCAGGAGC -3′	183	JQ359443
AvBD2 (real time RT-PCR)	5′-GATTGTCTTCGCCCCAGCGGGA -3′	5′-TTATACATCCCATGGCGATTTG-3′	137	HQ909024
AvBD3 (real time RT-PCR)	5′-GAACTGCCACTCAGTGCAGAAT -3′	5′-ATGGGGGTTGTTTCCAGGAGC-3′	182	NM_204650
AvBD4 (real time RT-PCR)	5′-ATCGTGCTCCTCTTTGTGGCAGTTCA-3′	5′-CTACAACCATCTACAGCAAGAATACT-3′	171	–
AvBD5 (real time RT-PCR)	5′-GCTGTCCCTTCCTCGAGGATC-3′	5′-GGAATACCATCGGCTCCGGC-3′	139	HM452159
AvBD6 (real time RT-PCR)	5′-GTCAGCCCTACTTTTCCAGC -3′	5′-GCCCACCTGTTCCTCACAC -3′	142	JQ359442
AvBD7 (real time RT-PCR)	5′-CCAAGGAGTGGCAGGTCAG -3′	5′-CCTCGCACAGCAGGAACC -3′	142	KR018386
AvBD9 (real time RT-PCR)	5′-GCTTACAGCCAAGGAGATGCT-3′	5′-GGAGCTAGGTGCCCATTTGCA-3′	144	HQ909023
AvBD10 (real time RT-PCR)	5′-GGTTCTGCCGATGCGCCTTTTG-3′	5′-CTGCGCCGGAATCTTGGCAC-3′	147	HQ909025
AvBD12 (real time RT-PCR)	5′-TCACGGGCACGCACACG-3′	5′-GAGAATGACGGGCTCGAAGC-3′	115	KR018387
AvBD16 (real time RT-PCR)	5′-TGGAATCAGAGGTGGAGTTTG-3′	5′-GCAGCACGGGACGATAAG-3′	88	–
TLR1 (real time RT-PCR)	5′-AGTCCATCTTTGTGTTGTCGCC-3′	5′-ATTGGCTCCAGCAAGATCAGG-3′	127	JF823982
TLR2 (real time RT-PCR)	5′-GATTGTGGATAACATCATCGACTC-3′	5′-AGAGCTGCTTTCAAGTTTTCCC-3′	294	JN982474.1
TLR3 (real time RT-PCR)	5′-TCAGTACATTTGCAACACTC-3′	5′-AGCATCATAATCATACCCTTCC-3′	256	KC292270.1
TLR4 (real time RT-PCR)	5′-AATCTGAAATCACTGAGCTCAAAT-3′	5′-GAGATGTTAAACCATAGAAG-3′	190	HQ436371.1
TLR5 (real time RT-PCR)	5′-CCTTGTGCTTTGAGGAACGAGA-3′	5′-CACCCATCTTTGAGAAACTGTC-3′	124	JQ711137.1
TLR7 (real time RT-PCR)	5′-TTCTGGCCACAGATGTGACG-3′	5′-CCTTCAACTTCGCAGTGCAG-3′	219	JQ910168.1
TLR15 (real time RT-PCR)	5′-GTTCTCTCTCCCAGTTTTGTAAACAGC-3′	5′-AACGTTCATTTGTTGTTTTTAGGAC-3′	262	JQ014619.1
iNOS (real time RT-PCR)	5′- GAACAGCCAGCTCATCCGATA -3′	5′- CCCAAGCTCAATGCACAACTT -3′	103	U34045
18s (real time RT-PCR)	5′-TCCCAGTAAGCGCGAGTCAT-3′	5′-ACGGGCGGTGTGTACAAAG-3′	64	AB064942
IL-1β(real time RT-PCR)	5′-CCGCAAAGTGAGGCTCAACA-3′	5′-GTAGCCTTTGATGCCCAGC-3′	100	JF505290
IL-2 (real time RT-PCR)	5′-ACCGAGAGCTGACCAACTTT-3′	5′-ATCACCCACACTAAGAGCAT-3′	177	AY392557.1
IL-6 (real time RT-PCR)	5′-AGATGGTGATAAATCCTGATGA-3′	5′-CGGTTTTCTCCATAAATGAAGT-3′	150	HQ436372
IL-8 (real time RT-PCR)	5′-ATGAACGGCAAACTTGGGGCT-3′	5′-GCCAGAATTGCCTTTACGATCAG-3′	278	AB213393
IL-18 (real time RT-PCR)	5′-TGAAATCTGGCAGCGGAATGAAC -3′	5′-TCCCATGTTCTTCCTCACAACA-3′	135	JF 505289
IFN-γ (real time RT-PCR)	5′-GCCACACATCAAAAACCTGTCT-3′	5′-GGAGACTGGCTCCTTTTCCTT-3′	207	EU030377
MHC class I (real time RT-PCR)	5′-AAGAAGGAGAAGGGCTACAA-3′	5′-GACATGCAGATGCAATTATGC-3′	207	AY387653.1
FASLG (real time RT-PCR)	5′-TAACAGGAAACCCCACACAGC-3′	5′-CCGGAAGAGCACATTGGAGTA-3′	149	NM_001031559
18S rRNA (real time RT-PCR)	5′-TCC CAG TAA GCG CGA GTC AT-3′	5′-ACG GGC GGT GTG TAC AAA G-3′	65	AB064942

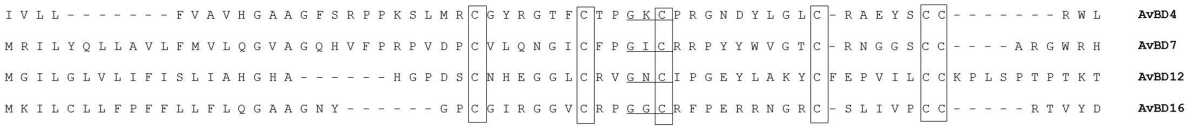


FIGURE 1 | Deduced amino acid sequence alignment of four novel avian β-defensin (AvBDs) from geese. The six conserved cysteines (C) are framed. The GXC motif is underlined. Dashes indicate that no identical or conserved residues were observed.

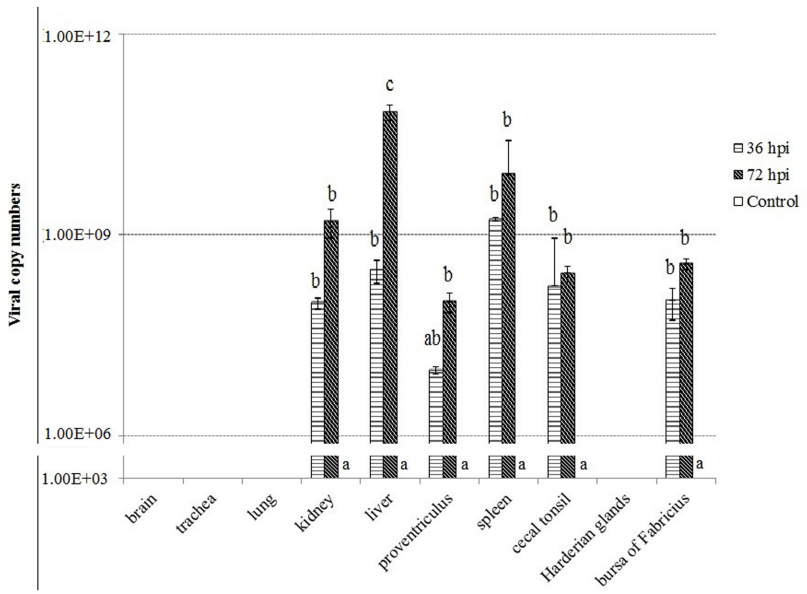


FIGURE 2 | Viral RNA copies in the tissues of geese in response to Newcastle disease virus (NDV) infection. Viral RNA copy numbers in the tissue samples from five geese in each group were measured by quantitative PCR at 36 and 72 hpi. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. [a,b,c]The values with different letters are significantly different ($P < 0.05$).

Other than the former four novel AvBDs, we tried to amplify other AvBDs (i.e., 8, 11, 13, and 14), which have not been characterized from geese till now. Unfortunately, none of these AvBDs were identified in the current study. In addition, partial sequences of 18S rRNA, TLRs (1, 2, 3, 5, 7, and 15), AvBDs (1–3, 5, 6, 9, and 10), cytokines (IL-1β, IL-2, IL-6, IL-8, IL-18, and IFN-γ), iNOS, MHC class I, and FASLG were amplified using respective primers (**Table 1**) from both spleen and bone marrow of healthy geese. The sequences of these peptides shared >95% identity to the respective sequences from geese that were available in a public database (data not shown).

Pathogenic Observation and Viral Replication in Geese

The geese in the control group did not show any clinical signs and none of these geese died during the experiment. In contrast, two geese died in group 2 on 3 dpi although none of the NDV-challenged geese showed clinical observations. Gross lesions, such as hemorrhage and edema in the proventriculus, hemorrhagic changes in the trachea, congestion in the lung, slight enlargement and congestion in livers were observed in the dead birds (Xu et al., accepted).

To understand the severity of pathology in the early infection period, we examined virus replication by real-time RT-PCR in ten tissues of geese of the control, 36 and 72 hpi, including brain, trachea, lungs, kidneys, liver, proventriculus, spleen, cecal tonsil, Harderian glands, and bursa of Fabricius. Samples from the control birds were negative. In contrast, the virus showed early replication, and was detected as early as 36 hpi, increased by 72 hpi in six tissues (i.e., kidneys, liver, proventriculus, spleen, cecal tonsil, and bursa of Fabricius). In addition, the virus replication was not detectable in brain, trachea, lung, and Harderian glands of infected geese (**Figure 2**).

Upregulation of TLR Expression in Response to NDV Infection

We analyzed expression patterns of TLR1–5, 7, and 15 of geese in response to NDV infection (**Figure 3**). The majority of these genes showed upregulation in infected geese. It is notable that TLR1 mRNA expression levels were upregulated significantly in several tissues of infected geese, including trachea by 72 hpi, lung by 36 hpi, cecal tonsil for the duration of the experiment, and Harderian glands by 36 hpi ($P < 0.05$). Despite lack of significant difference, expression of the other TLRs was also upregulated in

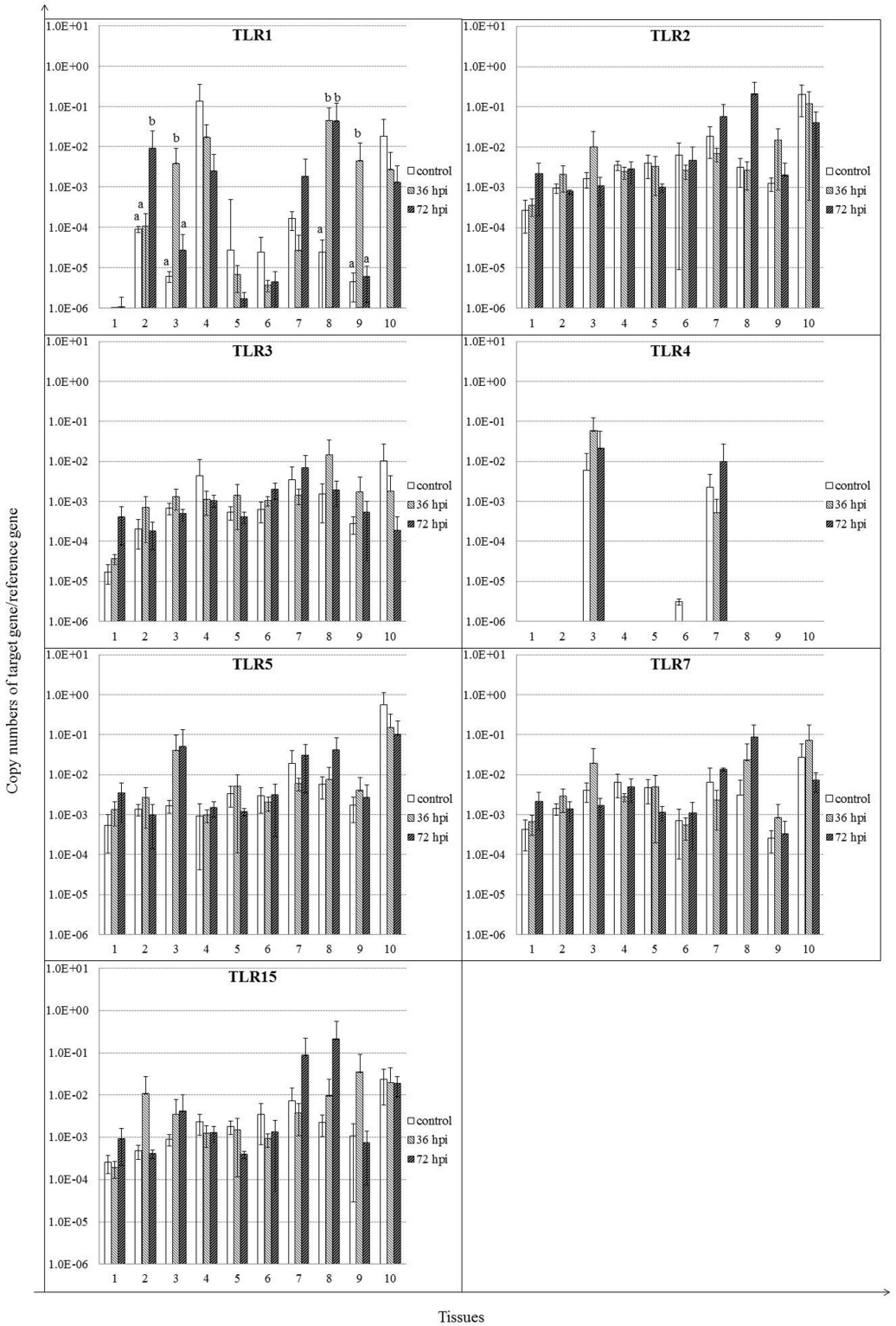


FIGURE 3 | Relative gene expression of toll-like receptors (TLRs) in the tissues of geese in response to NDV infection. (1) Brain, (2) Trachea, (3) lung, (4) kidneys, (5) Liver, (6) proventriculus, (7) spleen, (8) cecal tonsil, (9) Harderian gland, (10) bursa of Fabricius. cDNA copy numbers in the tissue samples from five geese of each group were measured by quantitative PCR at 36 and 72 hpi. TLR levels were normalized to the levels of 18S rRNA in the same samples. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. ^{a,b}The values with different letters are significantly different ($P < 0.05$).

most tissues of infected geese, including: TLR2 in cecal tonsil by72 hpi; TLR3 in brain by 72 hpi; TLR5 in lung by 36 and 72 hpi; TLR7 in cecal tonsil by 36 and 72 hpi; and TLR15 in trachea and Harderian glands by 36 hpi, and cecal tonsil by 72 hpi ($P > 0.05$). In contrast to extensive distribution of the above TLRs, TLR4 was detected only in lungs, proventriculus, and spleen, and no obvious difference was found between infected and control geese.

Differential Expression of AvBDs in NDV-Infected Geese

Most of the AvBDs measured were detectable in all the tissues from both the control and NDV-infected geese, except for AvBD1, 3 6, and 10 (**Figure 4**). AvBD3 was not detected in these tissues from either infected or control geese (data not shown). AvBD1 expression was completely suppressed in cecal tonsil by 36 and 72 hpi ($P < 0.05$). In the NDV-infected geese, there was significant upregulation in mRNA expression of AvBD5 in several tissues, including lung by 36 hpi, proventriculus by 72 hpi, and Harderian glands by 36 hpi, compared to the controls ($P < 0.05$). Expression of AvBD6, 7, 10, and 16 exhibited variable regulation in some tissues of geese in response to NDV infection. AvBD6 was significantly upregulated in trachea, lung, and cecal tonsil by 72 hpi, and in Harderian glands by 36 hpi ($P < 0.05$). In kidney, its expression was decreased by 36 hpi, and completely suppressed by 72 hpi ($P < 0.05$). AvBD7 was upregulated in trachea, but suppressed in liver for the duration of the experiment ($P < 0.05$). Furthermore, AvBD10 expression was upregulated in spleen by 72 hpi, but suppressed in Harderian glands by 72 hpi ($P < 0.05$). AvBD16 expression was increased by 36 hpi, but suppressed by 72 hpi in lung ($P < 0.05$). In addition, AvBD16 expression was upregulated in cecal tonsil for the duration of the experiment, and in Harderian glands by 36 hpi, compared to the controls ($P < 0.05$). In contrast to variable regulation of the above AvBDs in response to NDV infection, there was no significant regulation of AvBD2, AvBD4, and AvBD9 in all these tissues ($P > 0.05$).

NDV Infection Induces Cytokine Responses

We analyzed the expression patterns of IFN-γ and selected inflammatory cytokines (**Figure 5**). IFN-γ mRNA was induced by 36 hpi ($P > 0.05$), continued to increase by 72 hpi ($P < 0.05$) in both cecal tonsil and bursa of Fabricius. Expression of IL-8 was induced, although insignificantly by 36 hpi, and continued to rise by 72 hpi in spleen and cecal tonsil ($P < 0.05$). IL-8 mRNA was also induced in Harderian glands by 36 hpi ($P < 0.05$). Furthermore, IL-18 expression was increased significantly only in lung by 36 hpi ($P < 0.05$). For IL-1β, IL-2, and IL-6, little significant difference was detected at each time point in all tissues between the control and NDV-infected geese ($P > 0.05$).

Expression of iNOS in NDV-Infected Geese

Expression of iNOS was increased significantly only in trachea for the duration of the experiment ($P < 0.05$). In addition, despite the lack of significant differences, iNOS expression was also increased in lung and Harderian glands by 36 hpi ($P > 0.05$) (**Figure 6**).

Expression of MHC Class I and FASLG in Response to NDV Infection

MHC class I expression was induced, although insignificantly by 36 hpi ($P > 0.05$), and continued to rise by 72 hpi in proventriculus ($P < 0.05$). Furthermore, MHC class I was significantly induced in Harderian glands by 36 hpi ($P < 0.05$). In contrast, MHC class I expression was suppressed in bursa of Fabricius by 36 hpi ($P < 0.05$) and 72 hpi ($P > 0.05$). For FASLG, despite the lack of significant regulation, expression of FASLG was clearly induced in spleen by 72 hpi ($P > 0.05$) (**Figure 7**).

Relationship between Gene Expression of NDV and Immune Molecules in Spleen

Induction of host immune-related genes was accompanied by NDV replication in several tissues but their responses likely varied at each time point post-infection. Therefore, to clarify the relationship between viral replication and host immune response to NDV infection, we assessed the correlation between viral replication and gene expression in the spleen after infection, considering the importance of the spleen for both innate and adaptive immune responses (**Figure 8**). A significant positive correlation was shown between viral replication and mRNA expression for all TLRs measured by 72 hpi ($P < 0.05$), except for TLR15, but negative or few correlations by 36 hpi ($P > 0.05$). In addition, it is demonstrated that a high positive correlation was showed between viral replication and mRNA expression for TLR7 ($P < 0.05$). For AvBDs, significant positive correlation was observed only between viral replication and AvBD12 expression by 36 hpi ($P < 0.05$), whereas, we observed a significantly positive correlation between viral replication and expression of AvBD4, 5, 6, 10, and 12 by 72 hpi ($P < 0.05$) (**Figure 8**). Similarly, for cytokines, there was a high positive correlation between viral replication and mRNA expression for IFN-γ and IL-8 by 36 hpi ($P < 0.05$), while mRNA expression for all the cytokines had a high positive correlation with viral replication by 72 hpi ($P < 0.05$). Furthermore, mRNA expression of MHC class I and FASLG showed low correlation with viral replication by 36 hpi ($P > 0.05$), but high positive correlation with viral replication by 72 hpi ($P < 0.05$). iNOS expression showed a highly positive correlation with viral replication at both time points post-infection ($P < 0.05$). The results suggest that NDV infection caused an active host immune response mainly by 72 hpi.

DISCUSSION

In this study, viral RNA was detectable in kidneys, liver, proventriculus, spleen, cecal tonsil, Harderian glands, and bursa of Fabricius as early as 36 hpi, and increased at 72 hpi. Furthermore, high viral load was found in both liver and spleen of NDV-infected geese at each time point. In contrast, high level of viral replication was detected in all of these 10 tissues measured in dead geese (died on 3 dpi) (Xu et al., accepted). Surprisingly, no apparent respiratory signs were observed in all of the NDV-infected geese although 20%

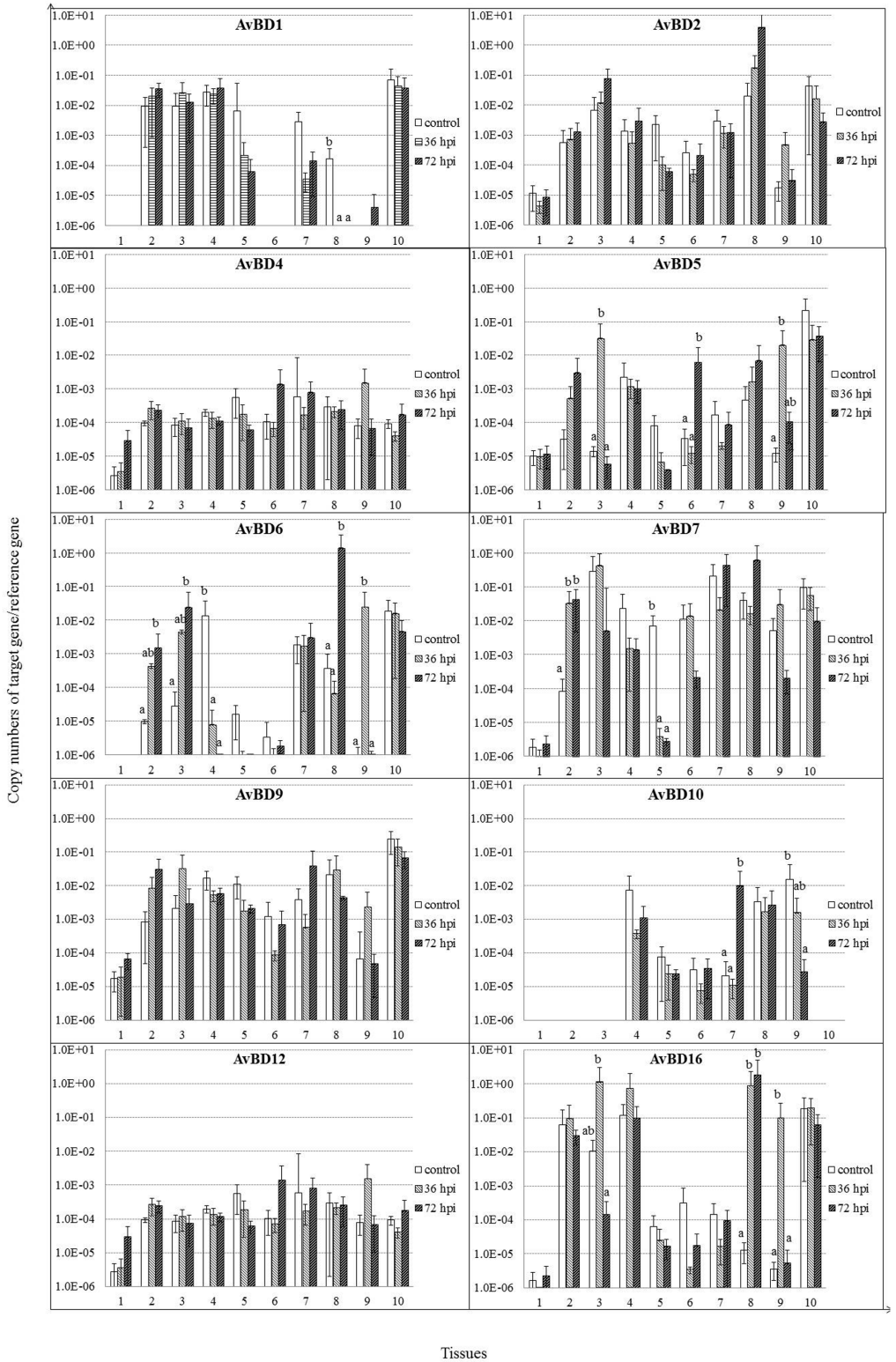


FIGURE 4 | Relative gene expression of AvBDs in the tissues of geese in response to NDV infection. (1) Brain, (2) Trachea, (3) lung, (4) kidneys, (5) Liver, (6) proventriculus, (7) spleen, (8) cecal tonsil, (9) Harderian gland, (10) bursa of Fabricius. cDNA copy numbers in the tissue samples from five geese of each group were measured by quantitative PCR at 36 and 72 hpi. AvBD levels were normalized to the levels of 18S rRNA in the same samples. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. ^{a,b}The values with different letters are significantly different ($P < 0.05$).

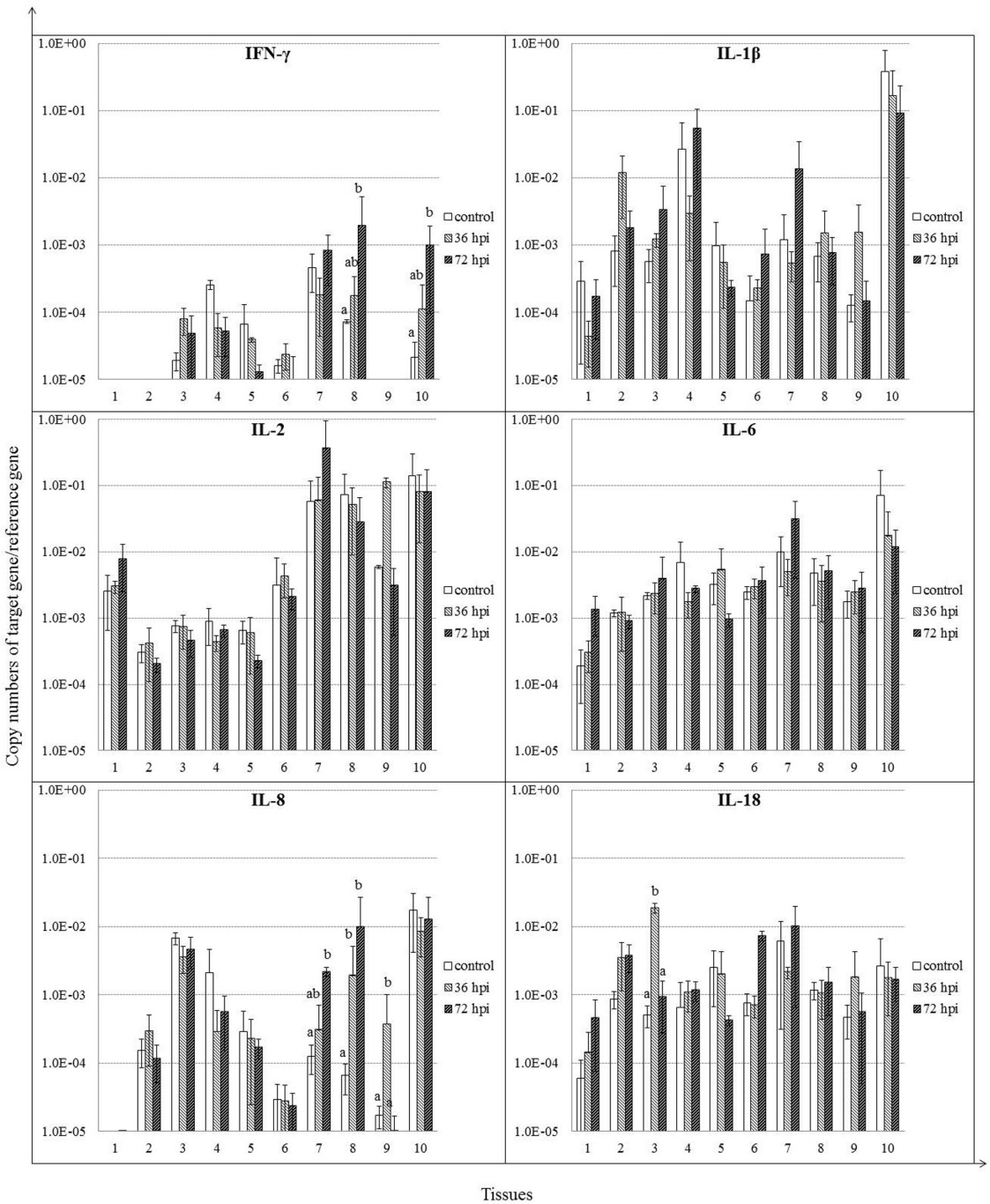


FIGURE 5 | Relative gene expression of cytokines in the tissues of geese in response to NDV infection. (1) Brain, (2) Trachea, (3) lung, (4) kidneys, (5) Liver, (6) proventriculus, (7) spleen, (8) cecal tonsil, (9) Harderian gland, (10) bursa of Fabricius. cDNA copy numbers in the tissue samples from five geese of each group were measured by quantitative PCR at 36 and 72 hpi. Cytokine levels were normalized to the levels of 18S rRNA in the same samples. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. [a,b]The values with different letters are significantly different ($P < 0.05$).

of the birds were died due to NDV infection (Xu et al., accepted). This was different from that of chickens which showed obvious respiratory signs after NDV infection (Ecco et al., 2011). In addition, gross lesions were also observed in various organs in dead goose, such as hemorrhage and edema in the proventriculus and slight enlargement and congestion in liver. Interestingly, no obvious gross lesions were observed in the trachea and lung of the live birds, in contrast to

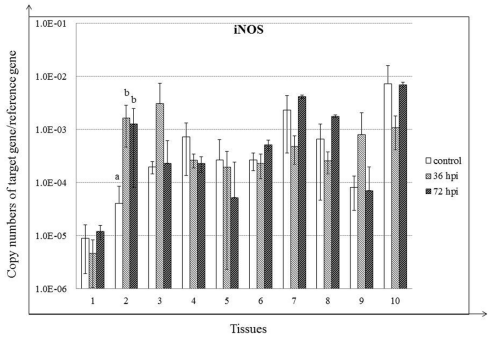


FIGURE 6 | Relative gene expression of iNOS in the tissues of geese in response to NDV infection. (1) Brain, (2) Trachea, (3) lung, (4) kidneys, (5) Liver, (6) proventriculus, (7) spleen, (8) cecal tonsil, (9) Harderian gland, (10) bursa of Fabricius. cDNA copy numbers in the tissue samples from five geese of each group were measured by quantitative PCR at 36 and 72 hpi. iNOS levels were normalized to the levels of 18S rRNA in the same samples. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. [a,b]The values with different letters are significantly different ($P < 0.05$).

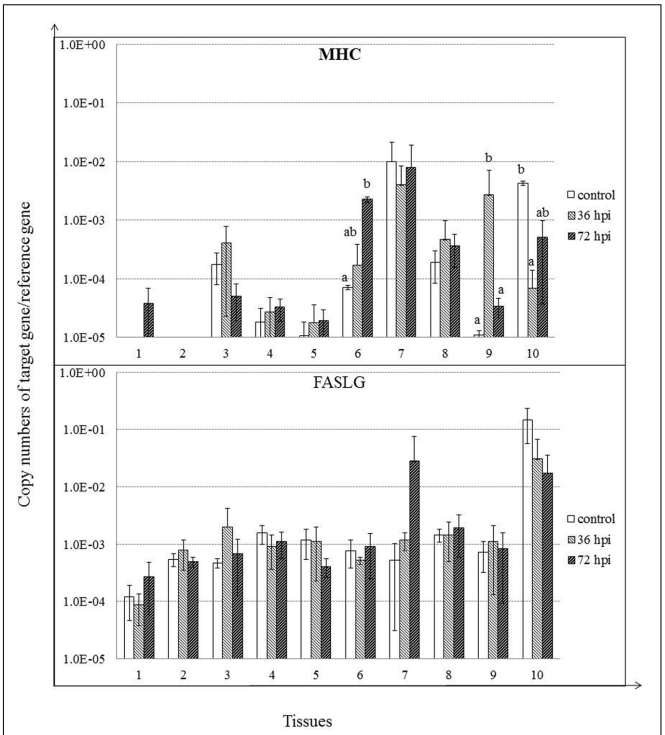


FIGURE 7 | Relative gene expression of MHC class I, and FASLG in the tissues of geese in response to NDV infection. (1) Brain, (2) Trachea, (3) lung, (4) kidneys, (5) Liver, (6) proventriculus, (7) spleen, (8) cecal tonsil, (9) Harderian gland, (10) bursa of Fabricius. cDNA copy numbers in the tissue samples from five geese of each group were measured by quantitative PCR at 36 and 72 hpi. Gene levels were normalized to the levels of 18S rRNA in the same samples. The control is the mean of results of Control-36 h and Control-72 h, due to results from both groups are almost the same. All assays were performed in triplicate, with five replicates per experiment, and each bar is the mean ± SD. [a,b]The values with different letters are significantly different ($P < 0.05$).

hemorrhagic changes in the trachea and congestion in the lung and brain of dead birds (Xu et al., accepted). These results suggested differences between chicken and goose after NDV infection.

The expression levels of 27 immune-related genes in 10 tissues of geese infected with the NDV strain were analyzed in the present study. The actual mechanisms responsible for host defense against viral replication are still not known. However, it is likely that TLRs play a part initially. TLRs have a role to activate the innate immunity by recognizing PAMPs, in mammals as well as birds (Hghihghi et al., 2010; Li et al., 2015). To date, TLRs have been identified in several avian species, such as duck, chicken, goose, and pigeon. Most TLRs have a potential role in antiviral responses, regardless of species (Hghihghi et al., 2010; Ma et al., 2012a, 2013; Li et al., 2015; Xu et al., 2015). In the current study, expression of TLR1–5, 7, and 15 was evaluated in NDV-infected geese, and most of these TLRs, except TLR4, were induced in different tissues. Consistent with the current results, recent evidence revealed that both TLR3 and TLR7 are induced by NDV in chickens, as well as by PPMV-1, a variant strain of NDV, in pigeons (Rue et al., 2011; Cheng et al., 2014; Li et al., 2015). It has also been demonstrated that overexpression of TLR3 enhances activity of IFN-β promoter and transcription factor nuclear factor-κB, thereby decreasing viral protein synthesis and titer (Cheng et al., 2014). These results strongly suggest that TLR1–3, 5, 7, and 15 actively participate in the recognition of the innate proinflammatory response after NDV infection.

An innate host response can be induced by the interaction between TLRs and their specific ligands, leading to the secretion of HDPs and cytokines (Abdel-Mageed et al., 2014). It is well established that defensins are the important components of host early innate immunity beyond cytokines (Ma et al., 2012a,b, 2013; Cuperus et al., 2013; Li et al., 2015; Xu et al., 2015). In poultry, only the β-defensins are present (Lynn et al., 2007). So far, more than 50 AvBDs have been characterized in different bird species (Cuperus et al., 2013). In recent studies, AvBD1–3, 5, 6, 9, and 10 have been isolated from geese (Ma et al., 2012b, 2013). Although initially described primarily as antibacterial agents, recent studies have also demonstrated direct antiviral potential (Ma et al., 2011, 2012a; Li et al., 2015; Xu et al., 2015). In addition to the former seven AvBDs, four novel AvBDs (4, 7, 12, and 16) were identified from geese in the present study. We found that expression of all of these AvBDs showed variable regulation in some tissues in geese in response to NDV infection. AvBD1 expression was suppressed at an early stage in cecal tonsil of geese. In contrast, AvBD1 expression was unchanged in tissues of chickens challenged with infectious bronchitis virus (Xu et al., 2015), or increased in tissues of ducks infected with duck hepatitis virus (Ma et al., 2012a). These findings suggest that effect of viruses on expression of AvBD1 depends upon the organs examined, the breed of birds, or viral strains. Furthermore, consistent with previous studies (Ma et al., 2011, 2012a), expressions of AvBD5 and AvBD12

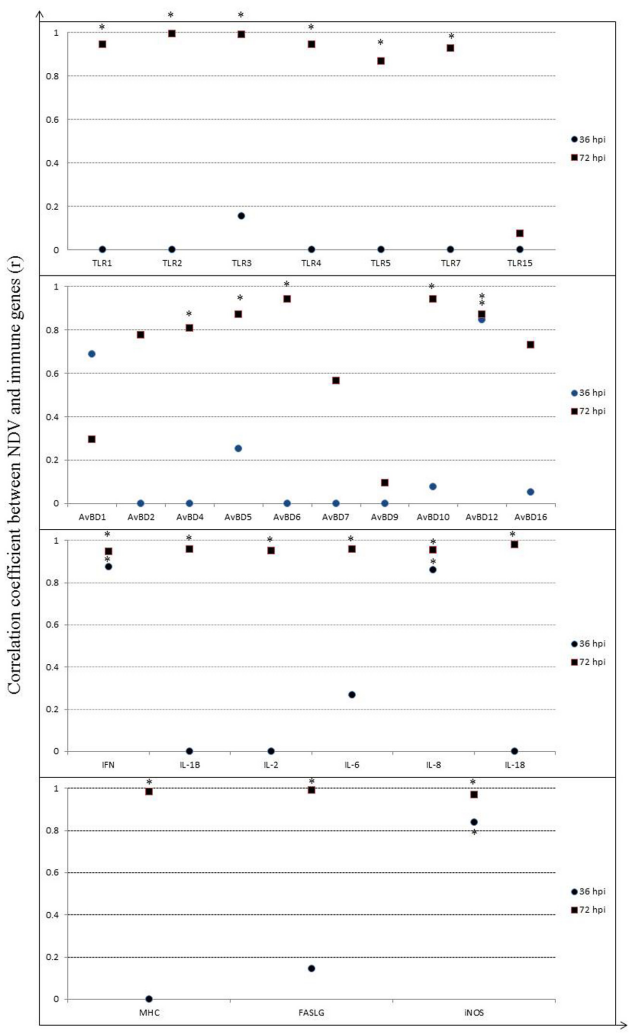


FIGURE 8 | Correlation between viral RNA copies and relative gene expression of host immune molecules in spleen of geese in response to NDV infection. Significant correlation is indicated by a star (*).

was upregulated in several tissues of geese in response to NDV infection. Regulation of AvBD6, 7, 10, and 16 expressions varied among tissues in geese in response to NDV challenge in this study. It was upregulated in some tissues but suppressed in others. These findings were partly consistent with previous studies on birds in response to other viral infections (Ma et al., 2011, 2012a; Li et al., 2015; Xu et al., 2015). To our surprise, a highly positive correlation was observed between viral replication and expressions of AvBD4–6, 9, 10, 12, and 16 in spleen at various time points post-infection; AvBDs have been shown to possess direct antiviral activity against viruses *in vitro* (Ma et al., 2011, 2012a; Li et al., 2015; Xu et al., 2015). However, the actual mechanisms responsible for this observation require further investigation.

In this study, we selected IFN-γ, IL-1β, IL-8, IL-2, IL-6, and IL-18 as indicators of antiviral and proinflammatory responses. These cytokines have been studied to understand the host immune response to a wide range of avian viruses and are important in infection control and virus clearance in birds (Ecco et al., 2011; Rue et al., 2011; Lee et al., 2013; Rasoli et al., 2014; Guan et al., 2015; Kang et al., 2015; Chimeno et al., 2016). The cytokine gene analysis showed that NDV infection in geese increased expressions of IFN-γ, IL-8, and IL-18. In addition, we also found the increased iNOS expression in NDV-infected geese. Moreover, the expression of these molecules highly correlates to the viral replication in the spleen. This result was similar to other reports which showed the correlation between high level of virus replication and intense inflammatory response caused by genotype VIId NDV in chickens and ducks (Ecco et al., 2011; Rue et al., 2011; Rasoli et al., 2014; Hu et al., 2015; Kang et al., 2015). The present findings are also in agreement with the observations in chickens infected by other avian viruses, including avian influenza virus (Lee et al., 2013; Guan et al., 2015), infectious bursal disease virus (Chimeno et al., 2016), and laryngotracheitis virus (Vagnozzi et al., 2016). Taken together, these results demonstrated that NDV infection induced strong innate immune responses and intense inflammatory responses at early stage in goose. These responses may associate with the viral pathogenesis.

Interestingly, MHC class I was significantly induced in proventriculus by 72 hpi and in Harderian glands by 36 hpi, whereas its expression was significantly suppressed in bursa of Fabricius by 36 hpi ($P < 0.05$). While, FASLG expression was slightly induced only in the kidneys and spleen, and remained at the basal level in the other tissues of geese in response to the NDV infection. The present result is inconsistent with that reported by Sarmento et al. (2008) in chickens infected by avian influenza H5N1 viruses. It is also found that expression of MHC class I was upregulated by avian influenza H3N2 virus (Tong et al., 2004). The possible reason might be that highly pathogenic viruses, such as avian influenza H5N1virus, have a mechanism to inhibit the expression of MHC class I. During infection with low pathogenic viruses such as avian influenza H3N2, as well as the current NDV strain, the viruses may be sensed by dendritic cells or macrophages and internalized into phagosomes. Here, they undergo proteolytic processing to produce antigenic peptides delivered by MHC-I molecules to the cell surface (Luo et al., 2014).

To our surprise, despite no viral replication detected in trachea and lung of infected geese at both time points, significant differences were found on expressions of several molecules, including TLRs (3, 5, and 15), AvBDs (5–7), IL-18, and iNOS in either of both tissues between the control and infected goose. Interestingly, result from our recent study showed viral RNA (go/CH/LHLJ/1/06) could be detected in both trachea and lung tissues in dead geese on 3 dpi (Xu et al., accepted). The result implies that host early defensing response against viral infection is active in both of trachea and lung, although the viral replication is later in both tissues than in the other tissues. However, the actual mechanisms responsible for the observation are not known and need to further study.

CONCLUSION

The present study demonstrated that NDV infection induces strong innate immune responses and intense inflammatory responses at early stage in goose which may associate with the viral pathogenesis. This study is the analysis on the host early immune response to challenge with virulent NDV and further investigation are required to characterize how NDV affects differential host responses of geese.

AUTHOR CONTRIBUTIONS

QX, YC, WZ, TQ, and TZ performed the experiments. CL performed the calculation. ZH and YS collected samples. DM and SL designed and conducted the study, and wrote the manuscript.

ACKNOWLEDGMENTS

The study was partly supported by Specialized Research Fund for the science and technological innovation talent of Harbin (2013RFXXJ019), Research Program for Applied Technology of Heilongjiang Province (PC13S02), Special Fund for Agro-scientific Research in the Public Interest (No. 201303033), and grants from the China Agriculture Research System (No. CARS-41-K12).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01587

REFERENCES

Abdel-Mageed, A. M., Isobe, N., and Yoshimura, Y. (2014). Effect of different TLR ligands on the expression of proinflammatory cytokines and avian beta-defensins in the uterine and vaginal tissues of laying hens. *Vet. Immunol. Immunopathol.* 162, 132–141. doi: 10.1016/j.vetimm.2014.10.013

Abdul-Careem, M. F., Haq, K., Shanmuganathan, S., Read, L. R., Schat, K. A., Heidari, M., et al. (2009). Induction of innate host responses in the lungs of chickens following infection with a very virulent strain of Marek's disease virus. *Virology* 393, 250–257. doi: 10.1016/j.virol.2009.08.001

Ahmed, K. A., Saxena, V. K., Ara, A., Singh, K. B., Sundaresan, N. R., Saxena, M., et al. (2007). Immune response to Newcastle disease virus in chicken lines divergently selected for cutaneous hypersensitivity. *Int. J. Immunogenet.* 34, 445–455. doi: 10.1111/j.1744-313X.2007.00722.x

Alexander, D. E. (1988). "Newcastle disease," in *A Laboratory Manual for the Isolation and Identification of Avian Pathogens*, eds H. G. Purchase, L. H. Arp, C. H. Domermuth, and J. E. Pearson (Kennett Sqare, PA: American Association for Avian Pathologists), 114–120.

Alexander, D. J. (2001). Gordon memorial lecture. Newcastle disease. *Br. Poult. Sci.* 42, 5–22.

Alexander, D. J., Morris, H. T., Pollitt, W. J., Sharpe, C. E., Eckford, R. L., Sainbury, R. M., et al. (1998). Newcastle disease outbreaks in domestic fowl and turkeys in Great Britain during 1997. *Vet. Rec.* 143, 209–212. doi: 10.1136/vr.143.8.209

Alexander, D. J., and Senne, D. A. (2008). "Newcastle disease, other avian paramyxoviruses, and pneumovirus infections," in *Diseases of Poultry*, 12th Edn, eds Y. M. Saif, A. M. Fadly, J. R. Glisson, L. R. McDougald, N. K. Nolan, and D. E. Swayne (Ames, IA: Blackwell Publishing), 75–115.

Ashkar, A. A., Yao, X. D., Gill, N., Sajic, D., Patrick, A. J., and Rosenthal, K. L. (2004). Toll-like receptor (TLR)-3, but not TLR4, agonist protects against genital herpes infection in the absence of inflammation seen with CpG DNA. *J. Infect. Dis.* 190, 1841–1849. doi: 10.1086/425079

Chen, F., Liu, J., Liu, D., Yan, Z., Ji, J., and Qin, J. (2012). Complete genome sequences of Newcastle disease virus strains isolated from broiler breeder flocks in China. *J. Virol.* 86, 12461–12462. doi: 10.1128/JVI.02228-12

Cheng, J. H., Sun, Y. J., Zhang, X. R., Zhang, F. Q., Zhang, S. L., Yu, S. Q., et al. (2014). Toll-like receptor 3 inhibits Newcastle disease virus replication through activation of pro-inflammatory cytokines and the type-1 interferon pathway. *Arch. Virol.* 159, 2937–2948. doi: 10.1007/s00705-014-2148-6

Chimeno, Z. S., Carballeda, J. M., Gravisaco, M. J., Lucero, M. S., Richetta, M., Gümez, E., et al. (2016). Immunomodulatory effect of baculovirus in chickens: how it modifies the immune response against infectious bursal disease virus. *Antiviral Res.* 131, 35–39. doi: 10.1016/j.antiviral.2016.04.004

Cuperus, T., Coorens, M., van Dijk, A., and Haagsman, H. P. (2013). Avian host defense peptides. *Dev. Comp. Immunol.* 41, 352–369. doi: 10.1016/j.dci.2013.04.019

Dai, Y., Liu, M., Cheng, X., Shen, X., Wei, Y., Zhou, S., et al. (2013). Infectivity and pathogenicity of Newcastle disease virus strains of different avian origin and different virulence for mallard ducklings. *Avian Dis.* 57, 8–14. doi: 10.1637/10298-070212-Reg.1

Ecco, R., Brown, C., Susta, L., Cagle, C., Cornax, I., Pantin-Jackwood, M., et al. (2011). In vivo transcriptional cytokine responses and association with clinical and pathological outcomes in chickens infected with different Newcastle disease virus isolates using formalin-fixed paraffin-embedded samples. *Vet. Immunol. Immunopathol.* 141, 221–229. doi: 10.1016/j.vetimm.2011.03.002

Guan, J., Fu, Q., and Sharif, S. (2015). Replication of an avian influenza virus and cytokine gene expression in chickens exposed by aerosol or intranasal routes. *Avian Dis.* 59, 263–268. doi: 10.1637/10972-110714-Reg

Guo, H. B., Liu, X. L., Xu, Y., Han, Z. X., Shao, Y. H., Kong, X. G., et al. (2014). A comparative study of pigeons and chickens experimentally infected with PPMV-1 to determine antigenic relationships between PPMV-1 and NDV strains. *Vet. Microbiol.* 168, 88–97. doi: 10.1016/j.vetmic.2013.11.002

Häuslaigner, R., Sonnenburg, J., Kothlow, S., Kaspers, B., Staubach, C., and Grund, C. (2009). Evaluation of an indirect enzyme-linked immunosorbent assay to study the specific humoral immune response of Muscovy ducks (*Cairina moschata*) and domestic geese (Anser anser var. domestica) after vaccination against Newcastle disease virus. *Avian Pathol.* 38, 89–95. doi: 10.1080/03079450902737813

Hghihghi, H. R., Read, L. R., Mohammadi, H., Pei, Y., Ursprung, C., Nagy, E., et al. (2010). Characterization of host responses against a recombinant fowlpox virus-vectored vaccine expressing the hemagglutinin antigen of an avian influenza virus. *Clin. Vaccine Immunol.* 17, 454–463. doi: 10.1128/CVI.00487-09

Hu, Z. L., Hu, J., Hu, S. L., Song, Q. Q., Ding, P. Y., Zhu, J., et al. (2015). High levels of virus replication and an intense inflammatory response contribute to the severe pathology in lymphoid tissues caused by Newcastle disease virus genotype VIId. *Arch. Virol.* 160, 639–648. doi: 10.1007/s00705-014-2301-2

Kang, Y., Li, Y., Yuan, R., Feng, M., Xiang, B., Sun, M., et al. (2015). Host innate immune responses of ducks infected with newcastle disease viruses of different pathogenicities. *Front. Microbiol.* 6:1283. doi: 10.3389/fmicb.2015.01283

Kumar, S., Nayak, B., Collins, P. L., and Samal, S. K. (2011). Evaluation of the Newcastle disease virus found in proteins in protective immunity by using a recombinant avian paramyxovirus type 3 vector in chickens. *J. Virol.* 85, 6521–6534. doi: 10.1128/JVI.00367-11

Lee, D., Yuk, S., Park, J., Kwon, J. M., Erdene-Ochir, T., Noh, J., et al. (2013). Innate immune response gene expression profiles in specific pathogen-free chickens infected with avian influenza virus subtype H9N2. *Biochip J.* 7, 393–398. doi: 10.1007/s13206-013-7411-5

Lee, E. K., Jeon, W. J., Kwon, J. H., Yang, C. B., and Choi, K. S. (2004). Molecular epidemiology of Newcastle disease viruses isolated in South Korea using sequencing of the fusion protein cleavage site region and phylogenetic relationships. *Avian Pathol.* 33, 482–491. doi: 10.1080/03079450400003700

Li, Y. Y., Xu, Q. Q., Zhang, T. T., Gao, M. Y., Wang, Q. L., Han, Z. X., et al. (2015). Host avian beta-defensin and toll-like receptor responses of pigeons following infection with pigeon Paramyxovirus type 1 virus. *Appl. Environ. Microbiol.* 81, 6415–6424. doi: 10.1128/AEM.01413-15

Liu, M., Shen, X., Cheng, X., Li, J., and Dai, Y. (2015). Characterization and sequencing of a genotype VIId Newcastle disease virus isolated from laying ducks in Jiangsu, China. *Genome Announc.* 3, e1412–e1415. doi: 10.1128/genomeA.01412-15

Liu, M., Wei, Y. Y., Dai, Y. B., Cheng, X., Zhou, S., Pan, Z. M., et al. (2010). Isolation and preliminary identification of a virulent Newcastle disease virus isolate of duck origin. *Chin. J. Anim. Infect. Dis.* 18, 67–71.

Liu, X. F., Wan, H. Q., Ni, X. X., Wu, Y. T., and Liu, W. B. (2003). Pathotypical and genotypical characterization of strains of Newcastle disease virus isolated from outbreaks in chicken and goose flocks in some regions of China during 1985-2001. *Arch. Virol.* 148, 1387–1403.

Luo, J., Carrillo, J. A., Menendez, K. R., Tablante, N. L., and Song, J. (2014). Transcriptome analysis reveals an activation of major histocompatibility complex 1 and 2 pathways in chicken trachea immunized with infectious larngotracheitis virus vaccine. *Poult. Sci.* 93, 848–855. doi: 10.3382/ps.2013-03624

Lynn, D. J., Higgs, R., Lloyd, A. T., O'Farrelly, C., Hervé-Grépinet, V., Nys, Y., et al. (2007). Avian beta-defensin nomenclature: a community proposed update. *Immunol. Lett.* 110, 86–89. doi: 10.1016/j.imlet.2007.03.007

Ma, D. Y., Lin, L. J., Zhang, K. X., Han, Z. X., Shao, Y. H., Liu, X. L., et al. (2011). Three novel *Anas platyrhynchos* avian β-defensins, upregulated by duck hepatitis virus, with antibacterial and antiviral activities. *Mol. Immunol.* 49, 84–96. doi: 10.1016/j.molimm.2011.07.019

Ma, D. Y., Zhang, K. X., Zhang, M. Y., Liu, X. L., Han, Z. X., Shao, Y. H., et al. (2012a). Identification, expression and activity analyses of five novel duck beta-defensins. *PLoS ONE* 7:e47743. doi: 10.1371/journal.pone.0047743

Ma, D. Y., Zhou, C. Y., Zhang, M. Y., Han, Z. X., Shao, Y. H., and Liu, S. W. (2012b). Functional analysis and induction of four novel goose (*Anser cygnoides*) avian β-defensins in response to salmonella enteritidis infection. *Comp. Immunol. Microbiol. Infect. Dis.* 35, 197–207. doi: 10.1016/j.cimid.2012.01.006

Ma, D. Y., Zhang, M. Y., Zhang, K. X., Liu, X. L., Han, Z. X., Shao, Y. H., et al. (2013). Identification of three novel avian beta-defensins from goose and their significance in the pathogenesis of *Salmonella*. *Mol. Immunol.* 56, 521–529. doi: 10.1016/j.molimm.2013.05.227

Marina, O. C., and Hanson, R. P. (1987). Cellular and humoral response of in vivo-bursectomized chickens to experimental challenge with velogenic Newcastle disease virus. *Avian Dis.* 31, 293–301. doi: 10.2307/1590875

Mase, M., Inoue, T., and Imada, T. (2009). Genotyping of Newcastle Disease Viruses Isolated from 2001 to 2007 in Japan. *J. Vet. Med. Sci.* 71, 1101–1104. doi: 10.1292/jvms.71.1101

Paul, M. S., Brisbin, J. T., Abdul-Csreem, M. F., and Sharif, S. (2013). Immunostimulatory properties of Toll-like receptor ligands in chickens. *Vet. Immunol. Immunopathol.* 152, 191–199. doi: 10.1016/j.vetimm.2012.10.013

Rasoli, M., Yeap, S. K., Tan, S. W., Moeini, H., Ideris, A., Bejo, M. H., et al. (2014). Alteration in lymphocyte responses, cytokine and chemokine profiles in chickens in fected with genotype VII and VIII velogenic Newcastle disease virus. *Comp. Immunol. Microbiol. Infect. Dis.* 37, 11–21. doi: 10.1016/j.cimid.2013.10.003

Rue, C. A., Susta, L., Cornax, I., Brown, C. C., Kapczynski, D. R., Suarez, D. L., et al. (2011). Virulent Newcastle disease virus elicits a strong innate immune response in chickens. *J. Gen. Virol.* 92, 931–939. doi: 10.1099/vir.0.025486-0

Sang, Y., Ross, C. R., Rowland, R. R., and Blecha, F. (2008). Toll-like receptor 3 activation decreases porcine arterivirus infection. *Viral Immunol.* 21, 303–313. doi: 10.1089/vim.2008.0042

Sarmento, L., Afonso, C. L., Estevez, C., Wasilenko, J., and Pantin-Jackwood, M. (2008). Differential host expression in cells infected with highly pathogenic H5N1 avian influenza viruses. *Vet. Immunol. Immunopathol.* 125, 291–302. doi: 10.1016/j.vetimm.2008.05.021

SAS software (1996). *SAS User's Guide: Statistics*. Cary, NC: SAS Institute Inc.

Seth, R. B., Sun, L., and Chen, Z. L. (2006). Antiviral innate immunity pathways. *Cell Res.* 16, 141–147. doi: 10.1038/sj.cr.7310019

Sick, C., Schneider, K., Staeheli, P., and Weining, K. C. (2000). Novel chicken CXC and CC chemokines. *Cytokine* 12, 181–186. doi: 10.1006/cyto.1999.0543

Sick, C., Schultz, U., Münster, U., Meier, J., Kaspers, B., and Staeheli, P. (1998). Promoter structures and differential responses to viral and nonviral inducers of chicken type I interferon genes. *J. Biol. Chem.* 273, 9749–9754. doi: 10.1074/jbc.273.16.9749

Sun, Y. J., Ding, N., Ding, S. S., Yu, S. Q., Meng, C. H., Chen, H. J., et al. (2013). Goose RIG-I functions in innate immunity against Newcastle disease virus infections. *Mol. Immunol.* 53, 321–327. doi: 10.1016/j.molimm.2012.08.022

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell* 140, 805–820. doi: 10.1016/j.cell.2010.01.022

Tong, H. H., Long, J. P., Li, D., and DeMaria, T. F. (2004). Alteration of gene expression in human middle ear epithelial cells induced by influenza A virus and its implication for the pathogenesis of otitis media. *Microb. Pathog.* 37, 193–204. doi: 10.1016/j.micpath.2004.06.012

Vagnozzi, A., Riblet, S., Zavala, G., Ecco, R., Afonso, C. L., and Garcßa, M. (2016). Evaluation of the transcriptional status of host cytokines and viral genes in the trachea of vaccinated and non-vaccinated chickens after challenge with the infections laryngotracheitis virus. *Avian Pathol.* 45, 106–113. doi: 10.1080/03079457.2015.1126804

Wan, H., Chen, L., Wu, L., and Liu, X. (2004). Newcastle disease in geese: natural occurrence and experimental infection. *Avian Pathol.* 33, 216–221. doi: 10.1080/0307945042000195803

Wang, J. J., Liu, H. L., Liu, W., Zheng, D. X., Zhao, Y. L., Li, Y., et al. (2015). Genomic characterizations of six pigeon paramyxovirus type 1 viruses isolated from live bird markets in China during 2011 to 2013. *PLoS ONE* 10:e0124261. doi: 10.1371/journal.pone.0124261

Xu, Y., Zhang, T. T., Xu, Q. Q., Han, Z. X., Liang, S. L., Shao, Y. H., et al. (2015). Differential modulation of avian β-defensin and Toll-like receptor expression in chickens infected with infectious bronchitis virus. *Appl. Microbiol. Biotechnol.* 99, 9011–9024. doi: 10.1007/s00253-015-6786-8

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Xu, Chen, Zhao, Zhang, Liu, Qi, Han, Shao, Ma and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 March 2017
doi: 10.3389/fmicb.2017.00402

180-Nucleotide Duplication in the G Gene of *Human metapneumovirus* A2b Subgroup Strains Circulating in Yokohama City, Japan, since 2014

Miwako Saikusa[1], Chiharu Kawakami[1], Naganori Nao[2], Makoto Takeda[2], Shuzo Usuku[1], Tadayoshi Sasao[1], Kimiko Nishimoto[1] and Takahiro Toyozawa[3]*

[1] Yokohama City Institute of Public Health, Yokohama, Japan, [2] Department of Virology III, National Institute of Infectious Diseases, Musashimurayama, Japan, [3] Yokohama City Public Health Center, Yokohama, Japan

Human metapneumovirus (HMPV), a member of the family *Paramyxoviridae*, was first isolated in 2001. Seroepidemiological studies have shown that HMPV has been a major etiological agent of acute respiratory infections in humans for more than 50 years. Molecular epidemiological, genetic, and antigenetic evolutionary studies of HMPV will strengthen our understanding of the epidemic behavior of the virus and provide valuable insight for the control of HMPV and the development of vaccines and antiviral drugs against HMPV infection. In this study, the nucleotide sequence of and genetic variations in the G gene were analyzed in HMPV strains prevalent in Yokohama City, in the Kanto area, Japan, between January 2013 and June 2016. As a part of the National Epidemiological Surveillance of Infectious Diseases, Japan, 1308 clinical specimens (throat swabs, nasal swabs, nasal secretions, and nasal aspirate fluids) collected at 24 hospitals or clinics in Yokohama City were screened for 15 major respiratory viruses with a multiplex reverse transcription–PCR assay. HMPV was detected in 91 specimens, accounting for 7.0% of the total specimens, and the nucleotide sequences of the G genes of 84 HMPV strains were determined. Among these 84 strains, 6, 43, 10, and 25 strains were classified into subgroups A2a, A2b, B1, and B2, respectively. Approximately half the HMPV A2b subgroup strains detected since 2014 had a 180-nucleotide duplication (180nt-dup) in the G gene and clustered on a phylogenic tree with four classical 180nt-dup-lacking HMPV A2b strains prevalent between 2014 and 2015. The 180nt-dup causes a 60-amino-acid duplication (60aa-dup) in the G protein, creating 23–25 additional potential acceptor sites for O-linked sugars. Our data suggest that 180nt-dup occurred between 2011 and 2013 and that HMPV A2b strains with 180nt-dup ($A2b_{180nt-dup}$ HMPV) became major epidemic strains within 3 years. The detailed mechanism by which the $A2b_{180nt-dup}$ HMPV strains gained an advantage that allowed their efficient spread in the community and the effects of 60aa-dup on HMPV virulence must be clarified.

Keywords: ***Human metapneumovirus*, molecular epidemiology, G gene, duplication, surveillance**

OPEN ACCESS

Edited by:
Akio Adachi,
University of Tokushima, Japan

Reviewed by:
Masato Tsurudome,
Mie University, Japan
Laymyint Yoshida,
Institute of Tropical Medicine, Nagasaki University (NEKKEN), Japan

****Correspondence:***
Miwako Saikusa
mi00-saikusa@city.yokohama.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *06 December 2016*
Accepted: *27 February 2017*
Published: *14 March 2017*

Citation:
Saikusa M, Kawakami C, Nao N, Takeda M, Usuku S, Sasao T, Nishimoto K and Toyozawa T (2017) 180-Nucleotide Duplication in the G Gene of Human metapneumovirus A2b Subgroup Strains Circulating in Yokohama City, Japan, since 2014. Front. Microbiol. 8:402. doi: 10.3389/fmicb.2017.00402

INTRODUCTION

The aim of this study was to strengthen our understanding of the epidemic behavior of *Human metapneumovirus* (HMPV) by providing virological data on the distribution patterns of HMPV. HMPV was first isolated in 2001 from young children suffering acute respiratory infections (ARIs), and was classified in the subfamily *Pneumovirinae* in the family *Paramyxoviridae* (van den Hoogen et al., 2001). Seroepidemiological studies have shown that the virus has been circulating globally for more than 50 years (van den Hoogen et al., 2001). HMPV is a major cause of upper and lower ARIs in infants and children (Collins and Karron, 2013), and also causes severe ARIs in aged adults and patients with underlying diseases (Boivin et al., 2002; Stockton et al., 2002; Falsey et al., 2003). Most children experience their first infection with HMPV before 5 years of age, but this infection does not provide lifelong immunity, and reinfection occurs frequently (van den Hoogen et al., 2001; Ebihara et al., 2004). In Japan, epidemics of HMPV have occurred between January and June, and especially in March and April (Kaida et al., 2006; Mizuta et al., 2010, 2013; Nakamura et al., 2013).

HMPV has a non-segmented negative-strand RNA genome of ∼13 kb. The HMPV genome contains eight genes in the order: 3′-N–P–M–F–M2–SH–G–L-5′ (Collins and Karron, 2013). The virus has three types of transmembrane viral proteins in its envelope, the fusion (F) protein, small hydrophobic (SH) protein, and glycoprotein (G protein), which are encoded by the F, SH, and G genes, respectively (Collins and Karron, 2013). The F protein is responsible for viral attachment and membrane fusion and is essential for viral infectivity (Biacchesi et al., 2004, 2005). Cell-surface integrins and glycosaminoglycans play roles in viral attachment and membrane fusion, mediated by the F protein (Cseke et al., 2009; Chang et al., 2012; Cox et al., 2012, 2015). The SH protein has properties consistent with those of viroporins and modulates the viral fusogenic activity (Masante et al., 2014). The G proteins of some lineages of HMPV also bind to glycosaminoglycans and contribute to HMPV infection (Thammawat et al., 2008; Adamson et al., 2012, 2013). The short cytoplasmic domain of the G protein inhibits the RIG-I-dependent signaling pathways (Bao et al., 2008). Despite these roles, the G and SH proteins are not essential for viral infectivity, but function as virulence factors (Biacchesi et al., 2004, 2005). Therefore, the G protein can be targeted in the development of antiviral drugs. The F protein is highly immunogenic and induces protective immunity, whereas the G and SH proteins are poorly immunogenic (Skiadopoulos et al., 2006).

The G protein is the most variable of the HMPV proteins, and mutations predominantly accumulate in the extracellular domain (Peret et al., 2004). The G protein has multiple potential glycosylation sites, and glycosylation can modify its immunogenicity (Liu et al., 2007). An evolutionary analysis of HMPV suggested that selective pressure is exerted on the G protein by the host's adaptive immunity (Gaunt et al., 2011). HMPV is divided to two groups, A and B, based on variations in its nucleotide sequence and its reactivity to monoclonal antibodies (van den Hoogen et al., 2004). The G gene has the most variable nucleotide sequence, and each viral group is further divided to two subgroups, A1 and A2 in group A, and B1 and B2 in group B, based mainly on the variations in the G gene (Biacchesi et al., 2003; van den Hoogen et al., 2004). Further detailed analyses of HMPV strains have also suggested two clades in the A2 subgroup, A2a and A2b (Huck et al., 2006). These different subgroups of HMPV have been detected in varying proportions in different countries and regions. In this study, we identified the genetic variations in the G gene of the HMPV strains prevalent in Yokohama City, in the Kanto area, Japan, between January 2013 and June 2016. Our data demonstrate a 180-nucleotide (nt) duplication (180nt-dup) in the G gene of HMPV and suggest that 180nt-dup occurred between 2011 and 2013. The HMPV A2b strains containing 180nt-dup ($A2b_{180nt-dup}$ HMPV strains) became major epidemic strains within 3 years, possibly overwhelming the classical A2b HMPV strains.

MATERIALS AND METHODS

Clinical Samples and HMPV Detection

In Yokohama City between January 2013 and June 2016, 1308 clinical specimens (throat swabs, nasal swabs, nasal secretions, and nasal aspirate fluids) were collected from patients suffering upper or lower ARIs in 16 sentinel hospitals and clinics (eight pediatric clinics, four internal medicine clinics, and four hospitals) participating in the National Epidemiological Surveillance of Infectious Diseases (NESID), instituted by the Infectious Diseases Control Law in Japan, and in eight other medical institutions (one pediatric clinic and seven hospitals) (Anon, 2010). There were 113 hospitals and 2962 clinics in Yokohama City, and the population in the city was ∼3.7 million. Thus, the numbers of hospitals and clinics participated in this study were ∼9.7 and ∼0.4%, respectively, in the city. Before collecting the clinical specimens in which to analyze the viruses causing ARIs, the physicians at each medical institution obtained the informed consent of the patients or their guardians. The de-identified clinical specimens were sent to Yokohama City Institute of Public Health and subjected to multiplex RT–PCR with the Seeplex® RV15 OneStep ACE Detection kit (Seegene, Seoul, South Korea), which identifies 15 major respiratory viruses. The clinical specimens that were positive for HMPV were analyzed further.

RNA Extraction and RT–PCR

The RNAs were purified from clinical specimens (140 μl) using the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's instruction, and dissolved in 60 μl of distilled water. The RNA in 5 μl of the purified RNA solution was reverse transcribed, and the G gene was amplified from the cDNA with RT–PCR using the PrimeScript II High Fidelity One Step RT–PCR Kit (TaKaRa Bio, Otsu, Japan) and a pair of G-gene-specific primers, SH7 (5′ primer)

(van den Hoogen et al., 2004) and GR (3′ primer) (Banerjee et al., 2011). The reaction mixture was prepared in a 50 μl solution, and the reverse-transcription reaction was performed at 45°C for 10 min. The reverse transcriptase was inactivated at 94°C for 2 min. The G gene cDNA was amplified with PCR with 40 cycles of 10 s at 98°C, 15 s at 56°C, and 10 s at 68°C. The cDNA products were separated electrophoretically in 3% NuSieve™ 3:1 Agarose gel in 0.5% TBE buffer. When no G-gene-specific band was detected or was only barely detectable, 5 μl of the RT–PCR product was subjected to semi nested PCR. In this assay, the reaction mixture was prepared in a 50 μl solution, and the RT–PCR products were denatured at 94°C for 2 min. The G gene was then PCR amplified with 40 cycles of 10 s at 98°C, 15 s at 50°C, and 30 s at 68°C with Tks Gflex™ DNA Polymerase (TaKaRa Bio) and a pair of G-gene-specific primers, SH7 (5′ primer) and GNR2 (3′ primer). The sequence of the GNR2 primer was 5′-GGATTCATTAAGAGGATCCA TTG-3′. The products of the semi nested PCR were separated electrophoretically to detect the G gene.

Sequencing

The nucleotide sequences of the G gene of the HMPV strains were determined with direct sequencing. The amplified HMPV G genes were purified with the QIAquick PCR Purification Kit (Qiagen). The PCR products were subjected to cycle sequencing with the BigDye Terminator ver. 1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and the SH7, GR, and GNR2 primers. The products of the cycle sequencing reactions were purified through Centri-Sep Spin Columns (Thermo Fisher Scientific), and the nucleotide sequences were determined with a 3500 Genetic Analyzer (Applied Biosystems). The nucleotide sequence data for the HMPV strains were aligned and edited with the BioEdit software (ver. 7.2.5) (Hall, 1999).

Phylogenetic Analysis

A multiple-sequence alignment was constructed with the MAFFT software (ver. 7.304b) using the default settings (Katoh et al., 2002). Phylogenetic analyses were performed with the maximum likelihood method in the MEGA software (ver. 7.0.20) (Kumar et al., 2016), and the statistical significance of the tree topologies was tested with bootstrapping (50 replicates).

N- and O-Glycosylation Site Analysis

Potential acceptor sites for *N*- and *O*-linked sugars were predicted with the NetNGlyc 1.0[1] and NetOGlyc 3.1 programs[2], respectively.

Nucleotide Sequence Accession Numbers

The nucleotide sequence data reported in the present study were deposited in the DDBJ/EMBL/GenBank nucleotide sequence database under accession numbers LC192170–LC192253.

Estimating the Evolutionary Rate of the HMPV G Gene

The overall rate of evolutionary change (nucleotide substitutions per site per year) in the HMPV G gene was estimated with the BEAST 2 program[3], which uses a Bayesian Markov Chain Monte Carlo (MCMC) approach (BEAST 2: A Software Platform for Bayesian Evolutionary Analysis[4]). An aligned G gene sequence dataset from the 84 HMPV strains detected in Yokohama City from 2013 to 2016 was analyzed with the BEAST 2 program using a strict HKY model. The MCMC chain was run for a sufficient length of time to ensure convergence (all expected sample size values exceeded 200 with a 10% burn-in).

Ethics Statement

These analyses have been done as a part of the National Epidemiological Surveillance of Infectious Diseases, Japan (NESID) as stipulated under the Infectious Diseases Control Law, and before collecting the clinical specimens, physicians in each medical institution obtained the informed consent of the patients or their guardians. The ethics committee of Yokohama City Institute of Public Health approved this study.

Prediction of the Year in Which 180nt-dup Occurred in the HMPV Genome

The consensus nucleotide sequence was determined using both the first 180-nucleotide sequence (1st-180nts) and the second

[1]http://www.cbs.dtu.dk/services/NetNGlyc/
[2]http://www.cbs.dtu.dk/services/NetOGlyc/
[3]http://beast2.org/
[4]http://dx.doi.org/10.1371/journal.pcbi.1003537

TABLE 1 | Characteristics of patients[a] infected with *Human metapneumovirus* (HMPV) strains.

	A2a	A2b	$A2b_{180nt-dup}$	B1	B2
Age[b] (year)	4.8 (2.2–5.2)	3.3 (1.8–6.5)	1.8 (1.4–6.4)	2.1 (1.3–4.4)	2.6 (1.3–4.7)
Sex[c]					
Female	3	10	4	6	7
Male	3	14	8	3	15
Body temperature[d]	39.1 (1.3)	38.6 (2.3)	39.4 (0.7)	38.7 (1.2)	39.7 (0.7)

[a]*Of the total 84 patients, clinical information was available for 73 patients.*
[b]*Data are medians (interquartile ranges).*
[c]*Number of patients.*
[d]*Mean (standard deviation).*



FIGURE 1 | Phylogenetic tree of 574 *Human metapneumovirus* (HMPV) strains, constructed with the maximum likelihood method. A phylogenetic tree was constructed from the G gene sequences of 84 HMPV strains and all other available HMPV G gene sequences. In total, 574 G gene sequences of HMPV strains were included. The HMPV A2b subgroup strains containing a 180-nt duplication (180nt-dup) in the G gene ($A2b_{180nt-dup}$ HMPV) are shown in cyan boxes and the other HMPV strains detected in Yokohama City between January 2013 and June 2016 are shown in orange boxes.

TABLE 2 | Numbers of HMPV strains detected in Yokohama City between January 2013 and June 2016 and their subgroup classification.

	180nt-dup[a]	Year of virus detection				Total
Subgroup		2013	2014	2015	2016	
A2a	–	4	1	1	0	6
A2b	–	11	3	12	2	28
	+	0	3	10	2	15
B1	–	2	2	1	5	10
B2	–	14	6	0	5	25
Total		31	15	24	14	84

[a] *180-nt duplication.*

180-nucleotide sequence (2nd-180nts) in the 15 $A2b_{180nt-dup}$ HMPV strains (in total, 30 sequences of 180 nt) and was assumed to be the corresponding 180 nt of the hypothetical ancestral strain of the 15 $A2b_{180nt-dup}$ HMPV strains. The number of nucleotide substitutions in the 1st- and 2nd-180nts of each $A2b_{180nt-dup}$ HMPV strain was determined and compared with the corresponding 180 nt of the hypothetical ancestral strain, and the year in which 180nt-dup occurred in the HMPV genome was estimated based on the mean rate of nucleotide substitutions for the HMPV G gene.

RESULTS

Nucleotide Sequence and Phylogenetic Analyses

Human metapneumovirus was detected in 91 of 1308 specimens from ARI patients, accounting for 7.0% of the specimens. The basic and clinical characteristics of the patients infected with HMPV are summarized in **Table 1**. The entire nucleotide sequences of the open reading frames of the G proteins were determined for 84 of the 91 HMPV strains. These 84 HMPV strains had 73 unique G gene sequences, containing at least one nucleotide substitution, whereas the remaining 11 sequences were identical to other G gene sequences. A phylogenetic analysis was performed using the G gene sequences of the 84 HMPV strains and all the HMPV G gene sequences available at the National Center for Biotechnology Information (NCBI) nucleotide sequence database[5]. In total, 574 G gene sequences of HMPV strains were used to construct the phylogenetic tree (**Figure 1**). Among our 84 strains, 6, 43, 10, and 25 strains were classified in subgroups A2a, A2b, B1, and B2, respectively (**Figure 1**). No HMPV strain belonged to subgroup A1. These data are similar to previous findings in Japan (Matsuzaki et al., 2008; Mizuta et al., 2010; Toda et al., 2010; Omura et al., 2011; Nidaira et al., 2012; Nakamura et al., 2013). **Table 2** shows the numbers of HMPV strains detected between January 2013 and June 2016 in Yokohama City and their subgroup classification. Multiple subgroup strains were detected in each year. Subgroup A2b strains were detected every year, and approximately half the A2b subgroup strains detected in 2014,

[5]https://www.ncbi.nlm.nih.gov/nucleotide/

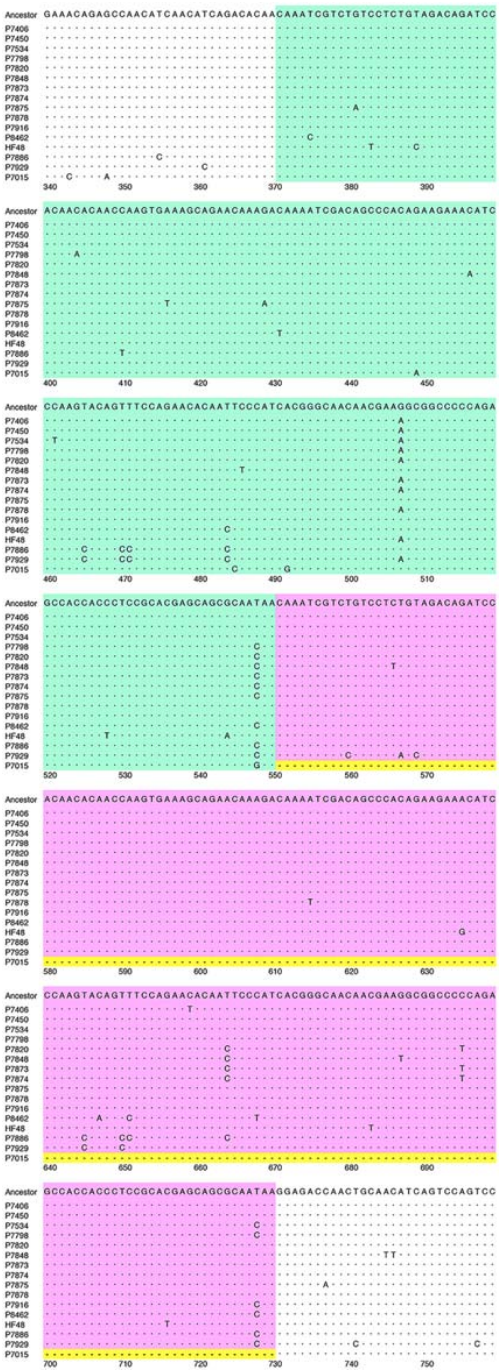


FIGURE 2 | Nucleotide sequence alignment of HMPV strains. An alignment of the plus-sense nucleotide sequences of the 180-nt duplicated regions and adjacent regions of the HMPV G genes of 15 $A2b_{180nt-dup}$ HMPV strains, the hypothetical ancestral strain (Ancestor), and a representative classical A2b strain (P7015) is shown. The first 180-nt region (1st-180nt) and the second 180-nt region (2nd-180nt) are shown in green and pink boxes, respectively. The first nucleotide of the initiation codon of the G gene is deemed to be nucleotide position 1. The nucleotide sequence of the hypothetical ancestral strain was predicted by constructing a consensus sequence of the duplicated 180-nt sequences of all 15 $A2b_{180nt-dup}$ HMPV strains. Dots indicate nucleotides identical to those of the hypothetical ancestral strain. Hyphens in yellow boxes indicate gaps in the sequence.

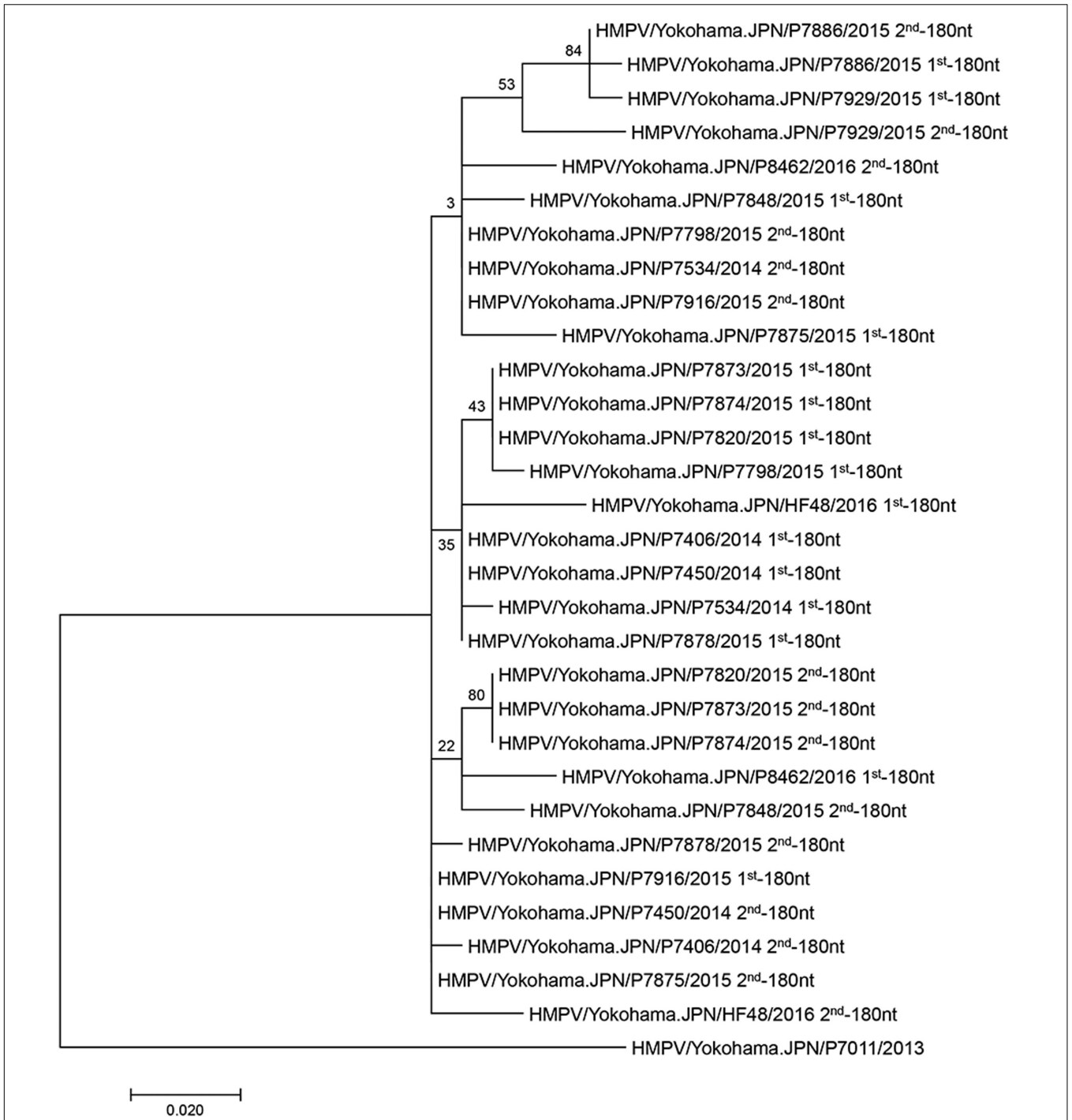


FIGURE 3 | Phylogenetic tree constructed with 1st-180nt and 2nd-180nt in the duplicated region of 15 A2b$_{180nt-dup}$ HMPV strains. The duplicated region contains the original 1st-180nt and the duplicated 2nd-180nt. A phylogenetic tree was constructed from the 1st- and 2nd-180nts of 15 A2b$_{180nt-dup}$ HMPV strains, and the tree was rooted with HMPV A2a strain HMPV/Yokohama.JPN/P7011/2013. The labels at the branch nodes indicate their statistical significance, calculated with bootstrapping (50 replicates).

2015, and 2016 contained 180nt-dup in the G gene (**Figure 2**). 180nt-dup is a duplication of the 180 nt at nucleotide positions 371–550 (the first nucleotide of the initiation codon of the G gene is deemed to be nucleotide position 1). No A2b strain detected in 2013 contained 180nt-dup. The 15 A2b$_{180nt-dup}$ HMPV strains formed a small cluster on the phylogenetic tree (**Figure 1**), and the cluster also contained four classical A2b strains, which lack 180nt-dup and were detected in 2014 and

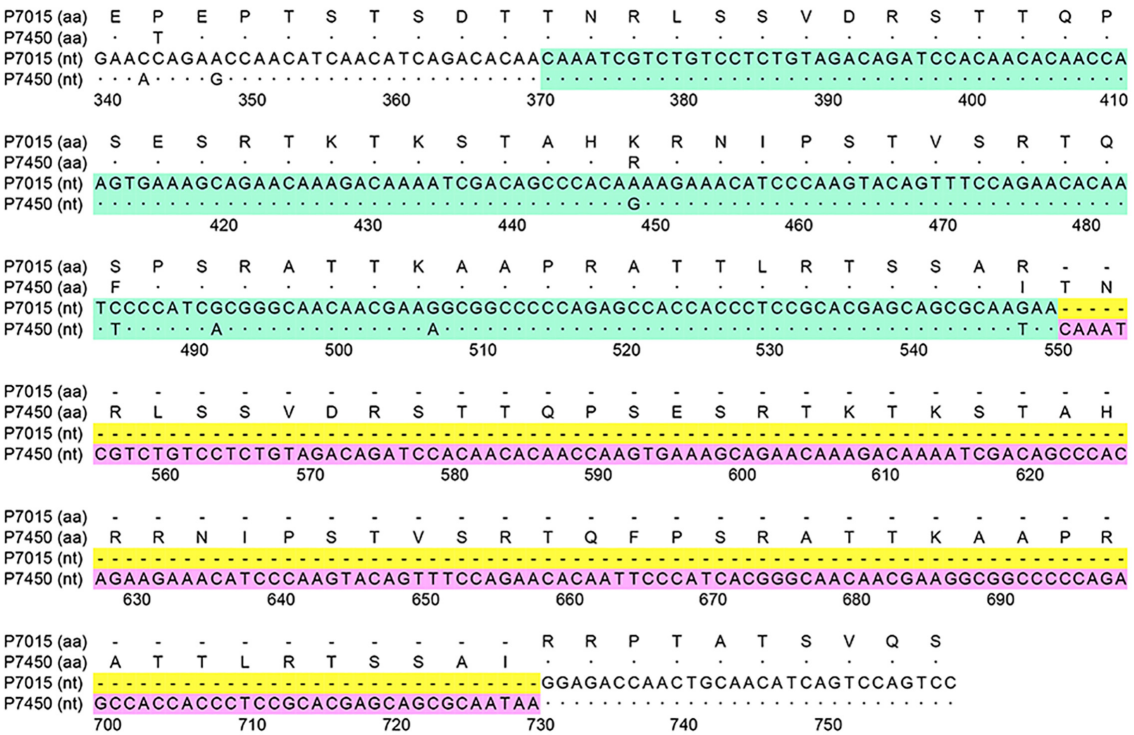


FIGURE 4 | Deduced amino acid sequence of A2b$_{180nt-dup}$ HMPV strain. A plus-sense nucleotide (nt) sequence alignment of a representative A2b$_{180nt-dup}$ HMPV strain (P7450) and a classical A2b HMPV strain (P7015), together with their deduced amino acid (aa) sequences, are shown. The 1st-180nt and 2nd-180nt regions are shown in green and pink boxes, respectively. The first nucleotide of the initiation codon of the G gene is deemed to be nucleotide position 1. Dots indicate nucleotides identical to those in the hypothetical ancestral strain. Hyphens in yellow boxes indicate gaps in the sequence.

2015 in Yokohama City. These data suggest that 180nt-dup occurred in a specific ancestor of the HMPV strains in this cluster (**Figure 2**).

Prediction of the Year in Which 180nt-dup Occurred

Although the 15 A2b$_{180nt-dup}$ HMPV strains seemed to be derived from a common ancestor, the 180nt-dup sequences of most A2b$_{180nt-dup}$ HMPV strains differed from one another (**Figure 2**). Even within the same viral genome, the 1st-180nt differed from the 2nd-180nt by one to six nucleotide substitutions. These observations were not unexpected, because substitutions would have occurred in each viral genome after the ancestral A2b strain had acquired 180nt-dup. RNA viruses are vulnerable to changes in their nucleotide sequences because the viral RNA-dependent RNA polymerase (RdRp) lacks proofreading activity. Therefore, the original 180-nt sequence in the hypothetical ancestral strain was predicted by constructing a consensus sequence of the 1st- and 2nd-180nts of the 15 A2b$_{180nt-dup}$ HMPV strains (**Figure 2**). The 360-nt sequence, consisting of the 1st- and 2nd-180nts of each of the 15 A2b$_{180nt-dup}$ HMPV strains, was then compared with that of the hypothetical ancestral strain (**Figure 2**). In the 1st-180nt, strain HMPV/Yokohama.JPN/P7916/2015 had the same nucleotide sequence as the ancestral strain, whereas the other 14 strains had one to six nucleotide substitutions relative to the sequence of the ancestral strain (**Figure 2**). In the 2nd-180nt, two strains (HMPV/Yokohama.JPN/P7450/2014 and HMPV/Yokohama.JPN/P7875/2015) had the same nucleotide sequence as the ancestral strain, whereas the other 13 strains had one to six nucleotide substitutions relative to the sequence of the ancestral strain (**Figure 2**). We then estimated the evolutionary rate of the HMPV G gene with the BEAST 2 program to predict the year in which 180nt-dup occurred. The evolutionary rate was estimated to be 4.3×10^{-3}/site/year (95% highest probability density: $3.2–5.4 \times 10^{-3}$/site/year), consistent with the rate previously estimated by de Graaf et al. (2008). Based on these data, 180nt-dup was predicted to have occurred in the ancestral strain between 2011 and 2013. **Figure 3** shows a phylogenetic tree constructed with the 15 1st-180nts and the 15 2nd-180nts. The tree was rooted with the HMPV A2a strain HMPV/Yokohama.JPN/P7011/2013. Theoretically, the 1st- and 2nd-180nts in each strain should be identical at the time when the duplication occurred. However, among the 15 A2b$_{180nt-dup}$ HMPV strains, the 1st- and 2nd-180nts of 13 strains have diverged randomly and the 1st- and 2nd-180nts of each of these 13 strains were not located in the same cluster. These data suggest that these 13 strains had the same ancestral strain but acquired nucleotide substitutions in both the 1st- and 2nd-180nts independently of each other. In contrast, the 1st- and 2nd-180nts of two A2b$_{180nt-dup}$ HMPV strains, HMPV/Yokohama.JPN/P7886/2015 and

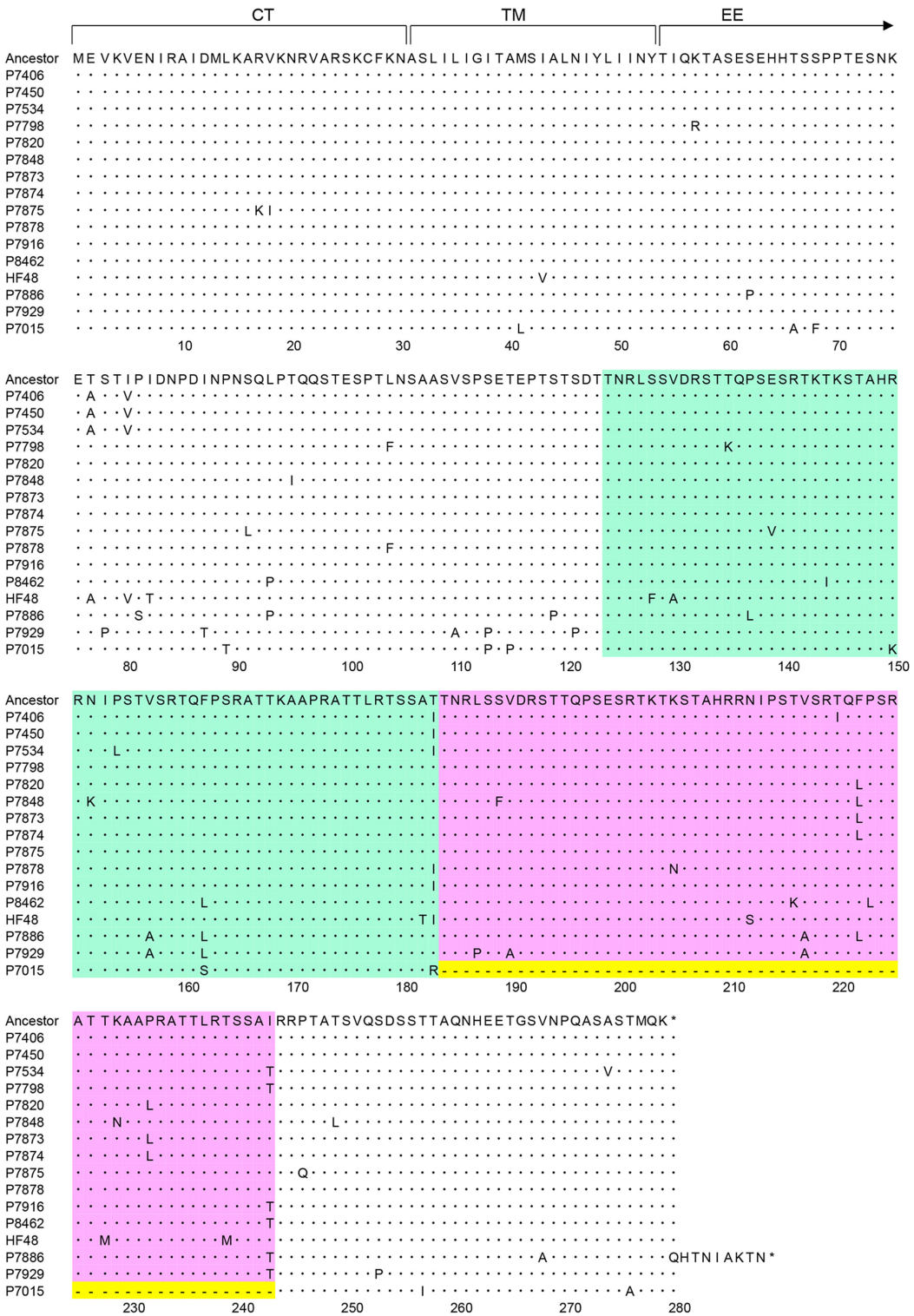


FIGURE 5 | Amino acid sequence alignment of HMPV strains. Aligned amino acid sequences derived from the HMPV G genes of 15 A2b$_{180nt-dup}$ HMPV strains, the hypothetical ancestral strain (Ancestor), and a representative classical A2b strain (P7015) are shown. The region between amino acid (aa) 1–30 is the cytoplasmic tail (CT) domain, and the region between aa 31–53 is the transmembrane (TM) domain. The remaining region is the extracellular ectodomain (EE). The 1st-180nt and 2nd-180nt regions are shown in green and pink boxes, respectively. The amino acid sequence of the hypothetical ancestral strain was predicted by constructing a consensus sequence of all 15 A2b$_{180nt-dup}$ HMPV strains. Dots indicate amino acids identical to those of the hypothetical ancestral strain. Hyphens in yellow boxes indicate gaps in the sequence.

HMPV/Yokohama.JPN/P7929/2015, were located in a distinct small cluster on the phylogenetic tree, suggesting that these two $A2b_{180nt-dup}$ HMPV strains had a different ancestral strain.

Deduced G-Protein Amino Acid Sequences of $A2b_{180nt-dup}$ HMPV Strains

Figure 4 shows a nucleotide sequence alignment of a representative $A2b_{180nt-dup}$ HMPV strain (HMPV/Yokohama.JPN/P7450/2014) and a classical A2b HMPV strain (HMPV/Yokohama.JPN/P7015/2013), and their deduced amino acid sequences. **Figure 5** shows amino acid sequence alignment of the G protein of 15 $A2b_{180nt-dup}$ HMPV strains, the hypothetical ancestral strain, and the representative classical A2b strain (HMPV/Yokohama.JPN/P7015/2013). Compared with the classical A2b strain, the $A2b_{180nt-dup}$ strains have a 60-amino-acid duplication (60aa-dup) in the G protein. Because the 180-nt region at positions 371–550 is duplicated, a codon (AGG) for arginine encoded at nucleotide positions 550–552 is replaced with a codon (ACA) for threonine (nucleotides in the duplicated sequence are underlined; the first nucleotide of the initiation codon of the G gene is deemed to be nucleotide position 1), which is followed by the duplicated nucleotide sequence encoding 60aa-dup (**Figure 4**). The 60aa-dup is located in the C-terminal half of the extracellular ectodomain (**Figure 5**) and contains many potential acceptor sites for O-linked sugars. The G protein of the classical A2b strains, which lack 60aa-dup, have 58–63 potential acceptor sites for O-linked sugars. With 60aa-dup, the G proteins of the $A2b_{180nt-dup}$ HMPV strains have acquired 23–25 additional potential acceptor sites for O-linked sugars. However, 60aa-dup does not affect the number of potential acceptor sites for N-linked sugars.

DISCUSSION

In the present study, we determined the G gene sequences of 84 HMPV strains detected in Yokohama City between January 2013 and June 2016. Our data demonstrate that HMPV strains of the four subgroups A2a, A2b, B1, and B2 were prevalent in Yokohama City in this period. Most importantly, the analysis detected 180nt-dup in the G genes of circulating HMPV A2b strains. Among the 32 A2b strains detected in Yokohama between 2014 and 2016, 15 (46.9%) contained 180nt-dup. Because 180nt-dup does not cause a frameshift in the G mRNA sequence, it generates the 60aa-dup in the G protein. A previous study (Leyrat et al., 2014) suggested that the extracellular ectodomain of the HMPV G protein sterically shields the F protein from recognition by host immune factors. These additional 60 amino acids may enhance this steric inhibition effect. The additional potential acceptor sites for O-linked sugars in the 60aa-dup region may also contribute to the evasion of immune recognition by the host, as has been observed for Ebola virus (Francica et al., 2010).

Similar duplications in the G gene have been reported in *Respiratory syncytial virus* (RSV), another member of the subfamily *Pneumovirinae* in the family *Paramyxoviridae*. A 72-nt duplication (72nt-dup) was detected in the G gene of genotype ON1 RSV strains in subgroup A (Eshaghi et al., 2012). A 60-nt duplication (60nt-dup) was also detected in the G gene of the genotype BA RSV strains in subgroup B (Trento et al., 2003). When we searched the currently available nucleotide sequence databases of viruses, we found no HMPV strain with a sequence duplication in the G gene. However, during the preparation of this manuscript, a similar report of 180nt-dup was presented at the 19th Annual Meeting of the European Society for Clinical Virology held in September 2016 (Pinana et al., 2016). Although only the abstract of the study is available, the study identified nine HMPV strains with 180nt-dup among 52 A2b strains detected in Barcelona, Spain, between 2014 and 2016 (Pinana et al., 2016). That study suggests that the virulence of the nine HMPV strains containing 180nt-dup was elevated based on the clinical manifestations of the children infected with those HMPV strains (Pinana et al., 2016). No such clinical differences were evident in our data.

Our data suggest that 180nt-dup occurred between 2011 and 2013. Importantly, these $A2b_{180nt-dup}$ HMPV strains became some of the major strains circulating in Yokohama City within 3 years. However, further surveillance data are required to determine whether these novel HMPV A2b strains persist as predominant strains. An increased frequency of $A2b_{180nt-dup}$ HMPV strains among the A2b strains was also observed in Barcelona, Spain (Pinana et al., 2016). Together, these data suggest that $A2b_{180nt-dup}$ HMPV strains are already circulating globally.

Although similar duplications in the G gene have been observed in RSV, the size of the duplication in the $A2b_{180nt-dup}$ HMPV strains is 2–3 times larger than that observed in the RSV G gene. Therefore, 60aa-dup, which results from 180nt-dup, may alter the viral antigenicity and the G protein functions more dramatically than the duplications observed in the RSV G gene. No significant difference in virulence was observed in the genotype ON1 and BA RSV strains when they acquired 72nt-dup and 60nt-dup, respectively, in their G genes (Sato et al., 2005; Tabatabai et al., 2014). However, these strains have spread rapidly and globally, and are currently the predominant strains in many countries (Trento et al., 2006; Duvvuri et al., 2015). Our data suggest that within 3 years, the $A2b_{180nt-dup}$ HMPV strains became one of the major strains of this pathogen. The mechanism by which the $A2b_{180nt-dup}$ HMPV strains gain an advantage over other strains when spreading in human populations, possibly by overwhelming the classical strains, and the effects of 180nt-dup in the G gene on HMPV virulence must be clarified.

AUTHOR CONTRIBUTIONS

MS, SU, TS, KN, and TT designed the study. MS, NN, and CK performed experiments. MS, NN, and MT analyzed data and wrote the paper. SU, TT, KN, and TS critically reviewed the manuscript.

FUNDING

This work was partly supported by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan and the Japan Foundation for Pediatric Research (No 15-004).

ACKNOWLEDGMENTS

We thank the staff of the clinics and hospitals that collected the specimens and clinical information. We are also grateful to all members of the Yokohama City Institute of Public Health for their technical support and dedicated assistance.

REFERENCES

Adamson, P., Thammawat, S., Muchondo, G., Sadlon, T., and Gordon, D. (2012). Diversity in glycosaminoglycan binding amongst hMPV G protein lineages. *Viruses* 4, 3785–3803. doi: 10.3390/v4123785

Adamson, P., Thammawat, S., Muchondo, G., Sadlon, T., Williams, J., and Gordon, D. (2013). hMPV lineage nomenclature and heparin binding. *Viruses* 5, 2546–2547. doi: 10.3390/v5102546

Anon (2010). Pathogen surveillance system in Japan and Infectious Agents Surveillance Report (IASR). *IASR* 31, 69–70.

Banerjee, S., Sullender, W. M., Choudekar, A., John, C., Tyagi, V., Fowler, K., et al. (2011). Detection and genetic diversity of human metapneumovirus in hospitalized children with acute respiratory infections in India. *J. Med. Virol.* 83, 1799–1810. doi: 10.1002/jmv.22176

Bao, X., Liu, T., Shan, Y., Li, K., Garofalo, R. P., and Casola, A. (2008). Human metapneumovirus glycoprotein G inhibits innate immune responses. *PLoS Pathog.* 4:e1000077. doi: 10.1371/journal.ppat.1000077

Biacchesi, S., Pham, Q. N., Skiadopoulos, M. H., Murphy, B. R., Collins, P. L., and Buchholz, U. J. (2005). Infection of nonhuman primates with recombinant human metapneumovirus lacking the SH, G, or M2-2 protein categorizes each as a nonessential accessory protein and identifies vaccine candidates. *J. Virol.* 79, 12608–12613. doi: 10.1128/JVI.79.19.12608-12613.2005

Biacchesi, S., Skiadopoulos, M. H., Boivin, G., Hanson, C. T., Murphy, B. R., Collins, P. L., et al. (2003). Genetic diversity between human metapneumovirus subgroups. *Virology* 315, 1–9. doi: 10.1016/S0042-6822(03)00528-2

Biacchesi, S., Skiadopoulos, M. H., Yang, L., Lamirande, E. W., Tran, K. C., Murphy, B. R., et al. (2004). Recombinant human Metapneumovirus lacking the small hydrophobic SH and/or attachment G glycoprotein: deletion of G yields a promising vaccine candidate. *J. Virol.* 78, 12877–12887. doi: 10.1128/JVI.78.23.12877-12887.2004

Boivin, G., Abed, Y., Pelletier, G., Ruel, L., Moisan, D., Cote, S., et al. (2002). Virological features and clinical manifestations associated with human metapneumovirus: a new paramyxovirus responsible for acute respiratory-tract infections in all age groups. *J. Infect. Dis.* 186, 1330–1334. doi: 10.1086/344319

Chang, A., Masante, C., Buchholz, U. J., and Dutch, R. E. (2012). Human metapneumovirus (HMPV) binding and infection are mediated by interactions between the HMPV fusion protein and heparan sulfate. *J. Virol.* 86, 3230–3243. doi: 10.1128/JVI.06706-11

Collins, P. L., and Karron, R. A. (2013). "Respiratory syncytial virus and metapneumovirus," in *Fields Virology*, eds D. M. Knipe, P. M. Howley, J. I. Cohen, D. E. Griffin, R. A. Lamb, M. A. Martin, et al. (Philadelphia, PA: Lippincott Williams & Wilkins), 1086–1123.

Cox, R. G., Livesay, S. B., Johnson, M., Ohi, M. D., and Williams, J. V. (2012). The human metapneumovirus fusion protein mediates entry via an interaction with RGD-binding integrins. *J. Virol.* 86, 12148–12160. doi: 10.1128/JVI.01133-12

Cox, R. G., Mainou, B. A., Johnson, M., Hastings, A. K., Schuster, J. E., Dermody, T. S., et al. (2015). Human metapneumovirus is capable of entering cells by fusion with endosomal membranes. *PLoS Pathog* 11:e1005303. doi: 10.1371/journal.ppat.1005303

Cseke, G., Maginnis, M. S., Cox, R. G., Tollefson, S. J., Podsiad, A. B., Wright, D. W., et al. (2009). Integrin alphavbeta1 promotes infection by human metapneumovirus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1566–1571. doi: 10.1073/pnas.0801433106

de Graaf, M., Osterhaus, A. D., Fouchier, R. A., and Holmes, E. C. (2008). Evolutionary dynamics of human and avian metapneumoviruses. *J. Gen. Virol.* 89(Pt 12), 2933–2942. doi: 10.1099/vir.0.2008/006957-0

Duvvuri, V. R., Granados, A., Rosenfeld, P., Bahl, J., Eshaghi, A., and Gubbay, J. B. (2015). Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: global and local transmission dynamics. *Sci Rep.* 5:14268. doi: 10.1038/srep14268

Ebihara, T., Endo, R., Kikuta, H., Ishiguro, N., Ishiko, H., Hara, M., et al. (2004). Human metapneumovirus infection in Japanese children. *J. Clin. Microbiol.* 42, 126–132. doi: 10.1128/JCM.42.1.126-132.2004

Eshaghi, A., Duvvuri, V. R., Lai, R., Nadarajah, J. T., Li, A., Patel, S. N., et al. (2012). Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. *PLoS ONE* 7:e32807. doi: 10.1371/journal.pone.0032807

Falsey, A. R., Erdman, D., Anderson, L. J., and Walsh, E. E. (2003). Human metapneumovirus infections in young and elderly adults. *J. Infect. Dis.* 187, 785–790. doi: 10.1086/367901

Francica, J. R., Verela-Rohena, A., Medvec, A., Plesa, G., Riley, J. L., and Bates, P. (2010). Steric shielding of surface epitopes and impaired immune recognition induced by the Ebola virus glycoprotein. *PLoS Pathog.* 6:e1001098. doi: 10.1371/journal.ppat.1001098

Gaunt, E. R., Jansen, R. R., Poovorawan, Y., Templeton, K. E., Toms, G. L., and Simmonds, P. (2011). Molecular epidemiology and evolution of human respiratory syncytial virus and human metapneumovirus. *PLoS ONE* 6:e17427. doi: 10.1371/journal.pone.0017427

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Huck, B., Scharf, G., Neumann-Haefelin, D., Puppe, W., Weigl, J., and Falcone, V. (2006). Novel human metapneumovirus sublineage. *Emerg. Infect. Dis.* 12, 147–150. doi: 10.3201/eid1201.050772

Kaida, A., Iritani, N., Kubo, H., Shiomi, M., Kohdera, U., and Murakami, T. (2006). Seasonal distribution and phylogenetic analysis of human metapneumovirus among children in Osaka City, Japan. *J. Clin. Virol.* 35, 394–399. doi: 10.1016/j.jcv.2005.12.009

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Leyrat, C., Paesen, G. C., Charleston, J., Renner, M., and Grimes, J. M. (2014). Structural insights into the human metapneumovirus glycoprotein ectodomain. *J. Virol.* 88, 11611–11616. doi: 10.1128/JVI.01726-14

Liu, L., Bastien, N., and Li, Y. (2007). Intracellular processing, glycosylation, and cell surface expression of human metapneumovirus attachment glycoprotein. *J. Virol.* 81, 13435–13443. doi: 10.1128/JVI.01469-07

Masante, C., El Najjar, F., Chang, A., Jones, A., Moncman, C. L., and Dutch, R. E. (2014). The human metapneumovirus small hydrophobic protein has properties consistent with those of a viroporin and can modulate viral fusogenic activity. *J. Virol.* 88, 6423–6433. doi: 10.1128/JVI.02848-13

Matsuzaki, Y., Itagaki, T., Abiko, C., Aoki, Y., Suto, A., and Mizuta, K. (2008). Clinical impact of human metapneumovirus genotypes and genotype-specific seroprevalence in Yamagata, Japan. *J. Med. Virol.* 80, 1084–1089. doi: 10.1002/jmv.21194

Mizuta, K., Abiko, C., Aoki, Y., Ikeda, T., Itagaki, T., Katsushima, N., et al. (2010). Endemicity of human metapneumovirus subgenogroups A2 and B2 in Yamagata, Japan, between 2004 and 2009. *Microbiol. Immunol.* 54, 634–638. doi: 10.1111/j.1348-0421.2010.00253.x

Mizuta, K., Abiko, C., Aoki, Y., Ikeda, T., Matsuzaki, Y., Itagaki, T., et al. (2013). Seasonal patterns of respiratory syncytial virus, influenza A virus, human metapneumovirus, and parainfluenza virus type 3 infections on the basis of virus isolation data between 2004 and 2011 in Yamagata, Japan. *Jpn. J. Infect. Dis.* 66, 140–145. doi: 10.7883/yoken.66.140

Nakamura, M., Hirano, E., Ishiguro, F., Mizuta, K., Noda, M., Tanaka, R., et al. (2013). Molecular epidemiology of human metapneumovirus from 2005 to 2011 in Fukui, Japan. *Jpn. J. Infect. Dis.* 66, 56–59. doi: 10.7883/yoken.66.56

Nidaira, M., Taira, K., Hamabata, H., Kawaki, T., Gushi, K., Mahoe, Y., et al. (2012). Molecular epidemiology of human metapneumovirus from 2009 to 2011 in Okinawa, Japan. *Jpn. J. Infect. Dis.* 65, 337–340. doi: 10.7883/yoken.65.337

Omura, T., Iizuka, S., Tabara, K., Tsukagoshi, H., Mizuta, K., Matsuda, S., et al. (2011). Detection of human metapneumovirus genomes during an outbreak of bronchitis and pneumonia in a geriatric care home in Shimane, Japan, in autumn 2009. *Jpn. J. Infect. Dis.* 64, 85–87.

Peret, T. C., Abed, Y., Anderson, L. J., Erdman, D. D., and Boivin, G. (2004). Sequence polymorphism of the predicted human metapneumovirus G glycoprotein. *J. Gen. Virol.* 85(Pt 3), 679–686. doi: 10.1099/vir.0.19504-0

Pinana, M., Vila, J., Gimferrer, L., Valls, M., Andrés, C., Ramón, J., et al. (2016). (Abstract no. 321 for Presentation at ESCV2016) Genetic variability of human metapneumovirus A strain circulating in Catalonia during the 2014–2015 and 2015–2016 seasons: a 180-nucleotide G gene duplication reported. *J. Clin. Virol.* 82S, S1–S142.

Sato, M., Saito, R., Sakai, T., Sano, Y., Nishikawa, M., Sasaki, A., et al. (2005). Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J. Clin. Microbiol.* 43, 36–40. doi: 10.1128/JCM.43.1.36-40.2005

Skiadopoulos, M. H., Biacchesi, S., Buchholz, U. J., Amaro-Carambot, E., Surman, S. R., Collins, P. L., et al. (2006). Individual contributions of the human metapneumovirus F, G, and SH surface glycoproteins to the induction of neutralizing antibodies and protective immunity. *Virology* 345, 492–501. doi: 10.1016/j.virol.2005.10.016

Stockton, J., Stephenson, I., Fleming, D., and Zambon, M. (2002). Human metapneumovirus as a cause of community-acquired respiratory illness. *Emerg. Infect. Dis.* 8, 897–901. doi: 10.3201/eid0809.020084

Tabatabai, J., Prifert, C., Pfeil, J., Grulich-Henn, J., and Schnitzler, P. (2014). Novel respiratory syncytial virus (RSV) genotype ON1 predominates in Germany during winter season 2012-13. *PLoS ONE* 9:e109191. doi: 10.1371/journal.pone.0109191

Thammawat, S., Sadlon, T. A., Hallsworth, P. G., and Gordon, D. L. (2008). Role of cellular glycosaminoglycans and charged regions of viral G protein in human metapneumovirus infection. *J. Virol.* 82, 11767–11774. doi: 10.1128/JVI.01208-08

Toda, S., Kimura, H., Noda, M., Mizuta, K., Matsumoto, T., Suzuki, E., et al. (2010). Phylogenetic analysis of human metapneumovirus from children with acute respiratory infection in Yamaguchi, Japan, during summer 2009. *Jpn. J. Infect. Dis.* 63, 139–140.

Trento, A., Galiano, M., Videla, C., Carballal, G., Garcia-Barreno, B., Melero, J. A., et al. (2003). Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J. Gen. Virol.* 84(Pt 11), 3115–3120. doi: 10.1099/vir.0.19357-0

Trento, A., Viegas, M., Galiano, M., Videla, C., Carballal, G., Mistchenko, A. S., et al. (2006). Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J. Virol.* 80, 975–984. doi: 10.1128/JVI.80.2.975-984.2006

van den Hoogen, B. G., de Jong, J. C., Groen, J., Kuiken, T., de Groot, R., Fouchier, R. A., et al. (2001). A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat. Med.* 7, 719–724. doi: 10.1038/89098

van den Hoogen, B. G., Herfst, S., Sprong, L., Cane, P. A., Forleo-Neto, E., de Swart, R. L., et al. (2004). Antigenic and genetic variability of human metapneumoviruses. *Emerg. Infect. Dis.* 10, 658–666. doi: 10.3201/eid1004.030393

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Saikusa, Kawakami, Nao, Takeda, Usuku, Sasao, Nishimoto and Toyozawa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 August 2016
doi: 10.3389/fmicb.2016.01263

Evolution and Transmission of Respiratory Syncytial Group A (RSV-A) Viruses in Guangdong, China 2008–2015

Lirong Zou[1†], Lina Yi[1,2†], Jie Wu[1], Yingchao Song[1], Guofeng Huang[1], Xin Zhang[1], Lijun Liang[1], Hanzhong Ni[1], Oliver G. Pybus[3], Changwen Ke[1] and Jing Lu[1,2,3]**

[1] Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China, [2] Guangdong Provincial Institute of Public Health, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China, [3] Department of Zoology, University of Oxford, Oxford, UK

OPEN ACCESS

Edited by:
Akio Adachi,
University of Tokushima, Japan

Reviewed by:
Hirokazu Kimura,
National Institute of Infectious Diseases, Japan
Charles Nyaigoti Agoti,
Kenya Medical Research Institute, Kenya
Lien Anh Ha Do,
Murdoch Childrens Research Institute, Australia

***Correspondence:**
Changwen Ke
kecw1965@aliyun.com
Jing Lu
jimlu0331@gmail.com

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 21 June 2016
Accepted: 02 August 2016
Published: 15 August 2016

Citation:
Zou L, Yi L, Wu J, Song Y, Huang G, Zhang X, Liang L, Ni H, Pybus OG, Ke C and Lu J (2016) Evolution and Transmission of Respiratory Syncytial Group A (RSV-A) Viruses in Guangdong, China 2008–2015. Front. Microbiol. 7:1263. doi: 10.3389/fmicb.2016.01263

Respiratory syncytial viruses (RSVs) including subgroups A (RSV-A) and B (RSV-B) are an important cause of acute respiratory tract infections worldwide. RSV-A include major epidemic strains. Fundamental questions concerning the evolution, persistence and transmission of RSV-A are critical for disease control and prevention, yet remain unanswered. In this study, we generated 64 complete *G* gene sequences of RSV-A strains collected between 2008 and 2015 in Guangdong, China. Phylogenetic analysis was undertaken by incorporating 572 publicly available RSV-A sequences. Current data indicate that genotypes GA1, GA4, and GA5 are endemic with limited epidemic activity. In contrast, the GA2 genotype which likely originated in 1980 has spread rapidly and caused epidemics worldwide. By analyzing GA2 genotype sequences across epidemic seasons within Guangdong, we find that RSV-A epidemics in Guangdong are caused by a combination of virus importation and local persistence, although the magnitude of the latter is likely overestimated due to infrequent sampling in other regions. Our results provide new insights into RSV-A evolution and transmission at global and local scales and highlights the rapid and wide spread of genotype GA2 compared to other genotypes. In order to control RSV transmission and outbreak, both local persistence and external introduction should be taken into account when designing optimal strategies.

Keywords: respiratory syncytial virus, phylogenetic, phylogeographic, evolution, transmission

INTRODUCTION

Human respiratory syncytial virus (RSV) is recognized as an important cause of acute respiratory tract infections (ARI), especially in children under 5 years old (Storey, 2010). The clinical manifestations of RSV infection range from mild symptoms in the upper respiratory tract to severe disease such as bronchiolitis and pneumonia (Welliver, 2003). RSV infection induces only partially protective immune responses that do not confer long-lasting protection (Gonzalez et al., 2012), therefore repeated infections of RSV are common (Henderson et al., 1979; Glezen et al., 1986). It has been estimated that RSV infects 70% of children during their first year of life and nearly all children older than two (Gonzalez et al., 2012). RSV infections cause an estimated 3.4 million children's hospital admissions annually, leading to a huge medical burden (Nair et al., 2010).

Respiratory syncytial viruse strains are classified into two major subgroups, RSV-A and RSV-B, according to their antigenic and genetic variability (Mufson et al., 1985). The two subgroups are further classified into different genotypes according to the genetic divergence of the viral *G* gene (Botosso et al., 2009). RSV-A include major epidemic strains (Scott et al., 2006). Based on *G* gene phylogenies, RSV-A can be classified into at least seven genotypes (GA1-7) (Trento et al., 2015). Co-circulation of different RSV-A genotypes in a population was previously considered as a reason for repeated infection and annual viral outbreaks (Garcia et al., 1994). However, recent surveillance has suggested that a single genotype of RSV-A, GA2, has spread internationally and become predominant in successive epidemic seasons (Eshaghi et al., 2012; Houspie et al., 2013; Agoti et al., 2014; Liu et al., 2014; Pierangeli et al., 2014; Duvvuri et al., 2015).

The prevention and control of RSV relies on our understanding of the virus' evolution and dissemination. Questions such as how RSV epidemics occur persist and reappear at global and local scales are important for designing optimal surveillance and prevention strategies, but remain largely unanswered (Hirano et al., 2014; Bose et al., 2015; Kimura et al., 2016). RSV is also proved as one of major etiologies of ARIs in mainland China (Liu et al., 2014; Dong et al., 2016; Fan et al., 2016). However, the dynamics of RSV infections in China are not clearly illustrated due to a lack of continuous surveillance on RSV epidemics. In this study, we collected RSV-A strains between 2008 and 2015 in Guangdong, China. These were sequenced and combined with other publicly available RSV-A sequences. We undertook phylogenetic, spatial and molecular clock analyses to investigate the molecular epidemiology of RSV-A at both global and local scales.

MATERIALS AND METHODS

Ethics Statement

In this study, all analyses were performed anonymously and did not involve human experimentation. This study was approved by the Ethics Review Committee of the Guangdong Center for Disease Control and Prevention. Respiratory samples were collected from patients in accordance with the guidelines of the Ministry of Health, P. R. of China, for public health purposes. Written consent was prepared and signed by all of patients or their guardian(s) when samples were collected.

Clinical Samples

Respiratory syncytial viruse surveillance was performed in four sentinel hospitals (Sun Yat-Sen Memorial Hospital, Guangdong No.2 Provincial People's Hospital,Guangzhou Women and Children's Medical Center and The Second Affiliated Hospital of Guangzhou Medical University) in Guangzhou, the capital city of Guangdong Province, from January 2008 to December 2015. Patients suspected of having ARIs from both inpatient and outpatient were enrolled according to these criteria: acute fever ($T \geq 38$°C), and/or abnormal leukocyte count, with any one respiratory symptom (such as sore throat, cough, expectoration, and dyspnoea/tachypnoea). Nasopharyngeal swabs (NPSs) were collected within 24 h after admission. The age and sex distributions of patient are shown in Supplementary Table S2.

Viral Test and RSV Sequencing

Nasopharyngeal swabs were kept and transported in viral transport medium and stored at −70°C prior to analysis. Total viral nucleic acids (DNA and RNA) were extracted using QIAamp MiniElute Virus Spin kits (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. For each specimen, RSV infection was detected by using qRT-PCR with a QIAGEN OneStep RT-PCR Kit and RSV *F* gene specific primers and a probe, RSV-F: 5′-GCGTAACWACACCTKTAAGCACT-3′ and RSV-R: 5′-CTTTGCTGYCTWACTATYTGAACATTG-3′ RSV-Probe: FAM-ATCAATGATATGCCTATAACAAATGA-BHQ1. PCR was performed with the following thermal profile: reverse transcription at 45°C for 10 min, followed by 10 min at 95°C; and 40 cycles of 95°C 15 s and 55°C 60 s. RSV-positive confirmed samples were further screened for subgroup (A/B) by amplifying and sequencing the full length of *G* genes which were further used for phylogenetic analysis. Briefly, RNA was firstly reverse transcribed into cDNA using random hexamer primers and SuperScript II Reverse Transcriptase (Invitrogen, USA). Then the *G* gene was amplified by using nested PCR with Taq PCR Master Mix Kit (Qiagen). The primers used for RSV-A are as follow: RSVA-F1 (5′-TCAAGCAAATTCTGGCCTTA-3′) and RSVA-R1 (5′-CAACTGCAATTCTGTTTACAGCA-3′); RSVA-F2 (5′-CCTTTGAGCTACCAAGAGCTC-3′) and RSVA-R2 (5′-GAGTGTGACTGCAGCAAGGA-3′). PCR was performed with the following thermal profile: 94°C 3 min, 40 cycles of 94°C 30 s, 52°C 60 s, and 72°C 1 min 40 s and followed by final extension at 72°C 10 min. Around 1300 bps PCR products were sequenced using an ABI3730xl DNA Analyzer at IGE Biotech Co., Ltd. (Guangzhou, China). The RSV-A sequences generated in this study have been submitted to GenBank (accession numbers KX009654- KX009717). Some RSV-A sequences collected from one sentinel hospital (Guangzhou Women and Children's Medical Center) between 2011 and 2013 were previously submitted and were included in analysis of virus local transmission (**Figure 3**).

Sequence Alignment and Maximum-Likelihood Phylogenetic Analysis

RSV-A sequences generated in this study were combined with all publicly available RSV-A *G* gene sequences with known sampling date and known sampling location in GenBank[1]. Partial sequences covering different parts of the *G* gene were excluded and identical sequences collected in the same sampling location on the same date were removed to improve computation time. In total, 572 sequences that covered at least 95% of *G* gene were included in the phylogenetic analysis (Supplementary Table S3). Multiple sequence alignment was performed using ClustalW (Larkin et al., 2007) and alignments were minimally edited by hand using Aliview (Larsson, 2014). Recombination was

[1]http://www.ncbi.nlm.nih.gov/genbank

checked by using the GARD tools (Kosakovsky Pond et al., 2006) available from the Datamonkey facility[2], which did not yield any indications of recombination being present in our data sets. The best-fit nucleotide substitution model (GTR + G) were selected by using W-IQ-TREE with the Bayesian information criterion (Trifinopoulos et al., 2016). Temporal accumulation of genetic divergence was assessed from maximum likelihood mid-point rooted phylogenies using the linear regression approach implemented in TempEst (formerly Path-O-Gen) (Rambaut et al., 2016).

Dated Phylogenetic Analysis

Bayesian Markov chain Monte Carlo (MCMC) phylogenetic inference was performed using BEAST, under a GTR + G substitution model (Shapiro et al., 2006) and a GMRF Bayesian skyride coalescent model (Minin et al., 2008). Preliminary analysis indicated high values for the coefficient of variation parameter of the molecular clock model, therefore an uncorrelated lognormal (UCLD) relaxed clock model was used in the final analysis to accommodate variation in substitution rates among branches (Drummond et al., 2006). Three independent MCMC runs of 1×10^8 steps were computed and 10–20% burn-in was discarded from each, resulting in a total of 2.0×10^8 total steps. Model parameters and trees were sampled every 10,000 MCMC steps. Convergence and behavior of MCMC chains was inspected using Tracer v1.6[3] (Lemey et al., 2014). A subset of 500 trees was randomly drawn from the combined posterior distribution of trees and used as an empirical distribution for subsequent phylogeographic analysis (Lemey et al., 2014). Maximum clade credibility (MCC) phylogenetic trees were also estimated for representative Guangdong RSV-A sequences and closely related sequences by setting strong priors value on virus evolution rates.

Phylogeography

We employed a Bayesian discrete phylogeographic approach to investigate viral spatial movement among four geographic regions (Africa, America, Asia, and Europe; **Figure 3**). Two RSV-A sequences from Australia were included in the Asia group. To ensure a realistic model of the direction of virus transmission, we used an asymmetric continuous-time Markov chain (CTMC) model (Edwards et al., 2011) to estimate ancestral locations and to estimate location posterior probabilities for each node in the time-scaled phylogenies.

RESULTS

Epidemiology of RSV in Guangdong

A total of 3843 ARI samples were collected from four surveillance hospitals in Guangzhou, the provincial capital of Guangdong, China, between 2008 and 2015. RSV was detected in 295 samples (7.68%, Supplementary Table S1). The seasonal distribution of RSV cases and the frequency of RSV positivity in ARI cases are shown in **Figure 1**. Although RSV infection peaks in each winter, from December to March, the seasonality of RSV infection in Guangdong is obscure, as a high relative risk of RSV infection is sometimes also observed in other seasons, e.g., in September 2011 and June 2015 (**Figure 1**). Notably, RSV epidemic activity was substantially increased in 2014 and 2015, with both more infection cases and higher positive rates in ARI samples in these years (**Figure 1**; Supplementary Table S1). RSV was detected in 14.1 and 13.8% of all tested samples collected in 2014 and 2015, which represents an almost threefold increase on the values for 2009–2013 (Supplementary Table S1). In total, 122 RSV positive samples were successfully sequenced in order to obtain full length *G* gene sequences; 64 samples were classified into the RSV-A subgroup through sequence alignment.

Genetic Evolution of RSV-A Through History

Spatial and temporal phylogenetic analyses were performed to describe the evolution of RSV-A. All other publicly available *G* gene sequences of RSV-A were combined with the new sequences generated in this study (see Materials and Methods). The evolutionary rate for RSV-A was estimated to be 2.3×10^{-3} substitutions/site/year (95% highest posterior density interval, HPD = $2.0–2.6 \times 10^{-3}$ substitutions/site/year). As **Figure 2** shows, the RSV-A *G* gene phylogeny was classified into major RSV genotypes, as previously described (Peret et al., 1998). Genotype GA1 represents older strains, mainly sampled between 1980 and 1990. Strains near the root of GA1 genotype primarily circulated in the USA (**Figure 2**). In addition, the spread of GA1 viruses appears to be geographically limited as most sequences in this genotype were identified in America. Although an American origin for the GA1 genotype seems the most probable (posterior probability = 1.0), this might be also caused by sampling biases (**Figure 2**). The GA1 genotype was rarely detected in the last 10 years, indicating that this genotype may be extinct or cause only sporadic infections. In comparison, the GA5 genotype has been continuously detected in the USA from 1979 to 2013 (**Figure 2**). Occasional GA5 infections have been also reported in countries on other continents, including Netherlands (Tan et al., 2012), Spain (Trento et al., 2015), Viet Nam (Do et al., 2015), and South Africa (Pretorius et al., 2013). Despite this, more than 90% of GA5 genotype strains were identified in American countries (**Figure 2**).

Genotypes GA2 and GA7 are closely related in the phylogenetic tree. The common ancestor of these two genotypes descended from an ancestral lineage that can be date to around 1978 (1970–1975, 95% HPD). GA7 viruses circulated for an only short period (1984–1998) and were rarely detected after 2000. GA2 viruses share a common ancestor around 1980 (1978–1982, 95% HPD) and have become the most prevalent genotype in the last decade. In contrast to the "endemic" pattern observed for genotypes GA1 and GA5, GA2 spread to countries outside the Americas after its emergence. Strains from different locations including Uruguay, the USA, Netherland and Korea contribute to the "trunk lineage" of the GA2 genotype (sequences collected 1990–2005, **Figure 2**). The geographic spread of RSV-A GA2

[2]http://www.datamonkey.org/

[3]tree.bio.ed.ac.uk

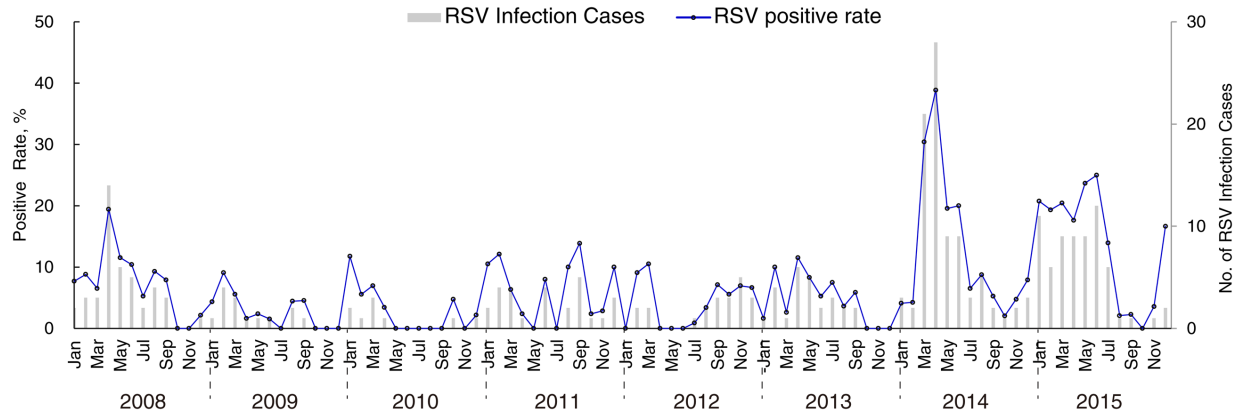


FIGURE 1 | Temporal distribution of RSV infections cases and RSV positive rates in acute respiratory tract infection samples collected in Guangzhou, Guangdong province, from 2008 to 2015.

was particularly pronounced after 2005 (**Figure 2**). After this, a significant change in RSV epidemiology was observed in different locations, with a shift from the circulation of multiple genotypes to prolonged circulation of predominant genotype GA2 (Botosso et al., 2009; Houspie et al., 2013). Most recently, the new variant of GA2, termed ON1, was identified in Ontario (Canada) and Panama in 2010 (Eshaghi et al., 2012). The ON1 genotype has spread widely in a short period of time, notably in 2011–2012, (**Figure 2**) and was reported as the dominant RSV-A strain in Europe in 2012–2013 (Pierangeli et al., 2014; Trento et al., 2015), in Africa in 2012 (Agoti et al., 2014) and in Asia in 2014 (Liu et al., 2014).

RSV-A Infection in Guangdong 2008–2015

The above analyses provided a basic overview on RSV-A evolution and transmission on a global scale. To further understand how RSV-A viruses circulated in a local area, we also estimated MCC phylogenetic trees from RSV-A sequences from Guangdong, China, together with closely related sequences from other regions. All Guangdong strains collected in 2008–2015 belong to genotype GA2. However, these strains were segregated into multiple subclusters and strains collected from the same epidemic season fell into different subclusters (**Figure 3A**). For example, strain 0263_GD-CHN_2008 is phylogenetically quite distinct from other strains collected in Guangdong 2008 such as 0185 and 0198, which grouped with contemporary strains from other countries.

In the phylogenetic tree (**Figure 3A**), we found several local clusters (GA2-2, GA2-3, and GA2-5) of Guangdong sequences, and also clusters that contained both Guangdong and non-Guangdong strains (GA2-1 and GA2-4). Local clusters contained viral strains collected from successive epidemic seasons, e.g., from 2011 to 2015 in the GA2-1 cluster. One interpretation of this result is that GA2 viruses have persisted in Guangdong between seasons. However, it is also equally possible that each local cluster was imported into Guangdong multiple times, but these importations are not observed due to the limited sampling of RSV from other locations. The presence of a sequence from Paraguay in cluster GA2-1 support the idea that there is substantial international movement of RSV lineages that is not being detected in current data due to limited virus sequence data from many regions. In the GA2-4 cluster, strain KM586822_GD-CHN_2011 is closely related to external strains sampled in the USA and India collected in previous years.

A similar pattern was observed for ON1 subgenotype viruses detected in Guangdong. The ON1 genotype is characterized by a 72 nt insertion in the viral *G* gene, resulting in 24 additional amino acids, of which 23 are duplications of amino acid positions 261–283 (Eshaghi et al., 2012). This new variant was first detected in Guangdong in 2012 (**Figure 3B**), and was predominant in Guangdong samples between 2014 and 2015, accounting for 9 of 9 strains sequenced in 2014 and for 36 of 38 (95%) strains analyzed in 2015 (**Figure 3B**). Local clusters of Guangdong strains collected between 2014 and 2015 were observed in the ON1 genotype (ON-1, ON-2, and ON-3, **Figure 3B**). In addition, Guangdong strains related to P14226_GD-CHN_2014 were more closely to a strain identified in Spain in 2012 (KF915233_SPA_2012) and this cluster is correspondingly termed ON-4. As discussed above, the amount of regional and international mixing of RSV observed in the phylogeny is likely underestimated due to limited sampling in many regions, although the highly similar sequences in cluster ON-3 probably do represent circulation within Guangdong or China itself, during the epidemics in 2014 and 2015. The amino acid alignment of the G protein of ON1 strains matches the corresponding phylogeny. **Figure 4** shows part of an alignment containing the variable mucin-like domains of the G protein (McLellan et al., 2013). The earliest ON1 strain from Guangdong (1119_GD-CHN_2012) shows 100% amino acid identity (across the full length of the G protein) to the first ON1 strain, identified in Canada in 2010. Viral strains within the clusters defined in **Figure 3B** show the same or highly similar amino acid changes. Several substitutions are specific to Guangdong and Chinese RSV strains belong to the ON1 genotype, specifically Lys216Asn, Ser299Arg and Pro300Ser.

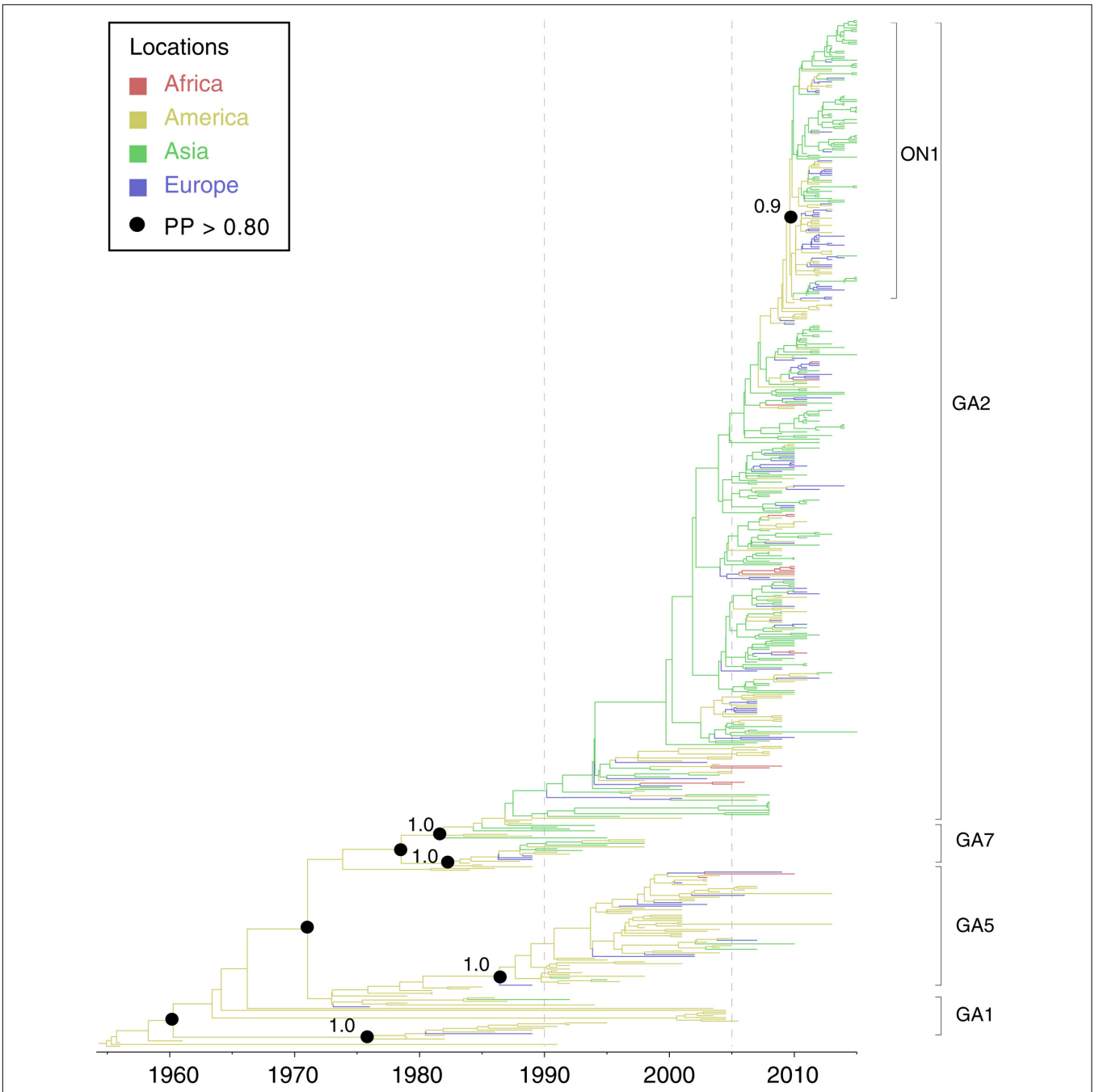


FIGURE 2 | Molecular clock phylogeny of global RSV-A *G* gene sequences. The tree shown is a maximum clade credibility (MCC) phylogeny, Branch colors represent the most probable ancestral locations of each branch, inferred using a phylogeographic model (see Materials and Methods for details). Four major genotypes of RSV-A are denoted on the right. Black circles indicate posterior probabilities >0.80 at selected nodes. Location posterior probabilities are noted above specific branches. Transverse axis shows time line in units of years.

DISCUSSION

RSV is one of the most important respiratory pathogen worldwide. Compared to human influenza viruses, the molecular epidemiology of RSV is largely unknown at both local and global scales. In this study, we undertook phylogenetic, spatial and molecular clock analyses on RSV-A by using sequences data from public database and from the surveillance in Guangdong China between 2008 and 2015. To achieve robust results in molecular clock and phylogeographic analyses, we used complete or nearly complete (>95%) *G* gene sequences in this study. RSV-A strains with partial sequences or without

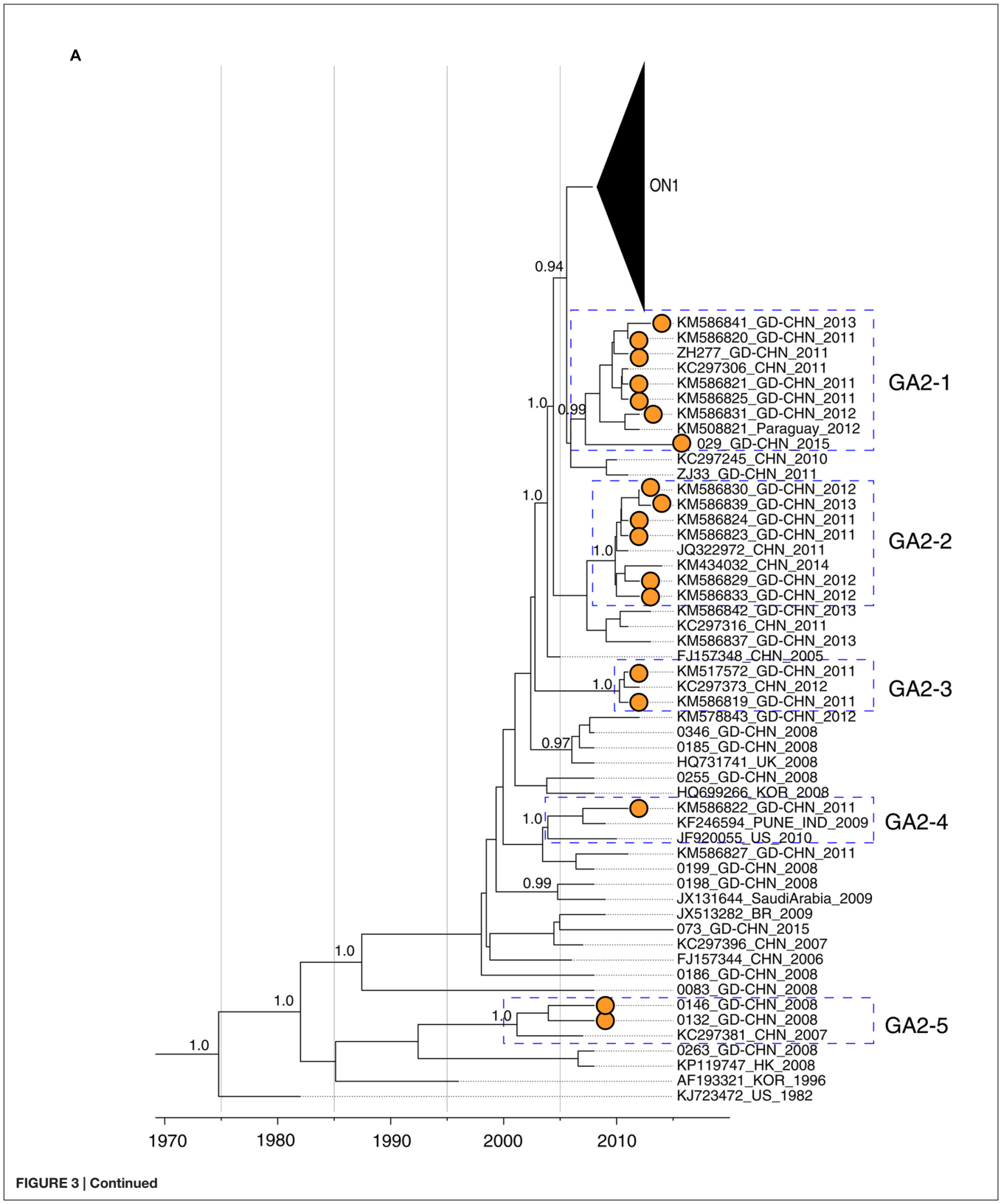


FIGURE 3 | Continued

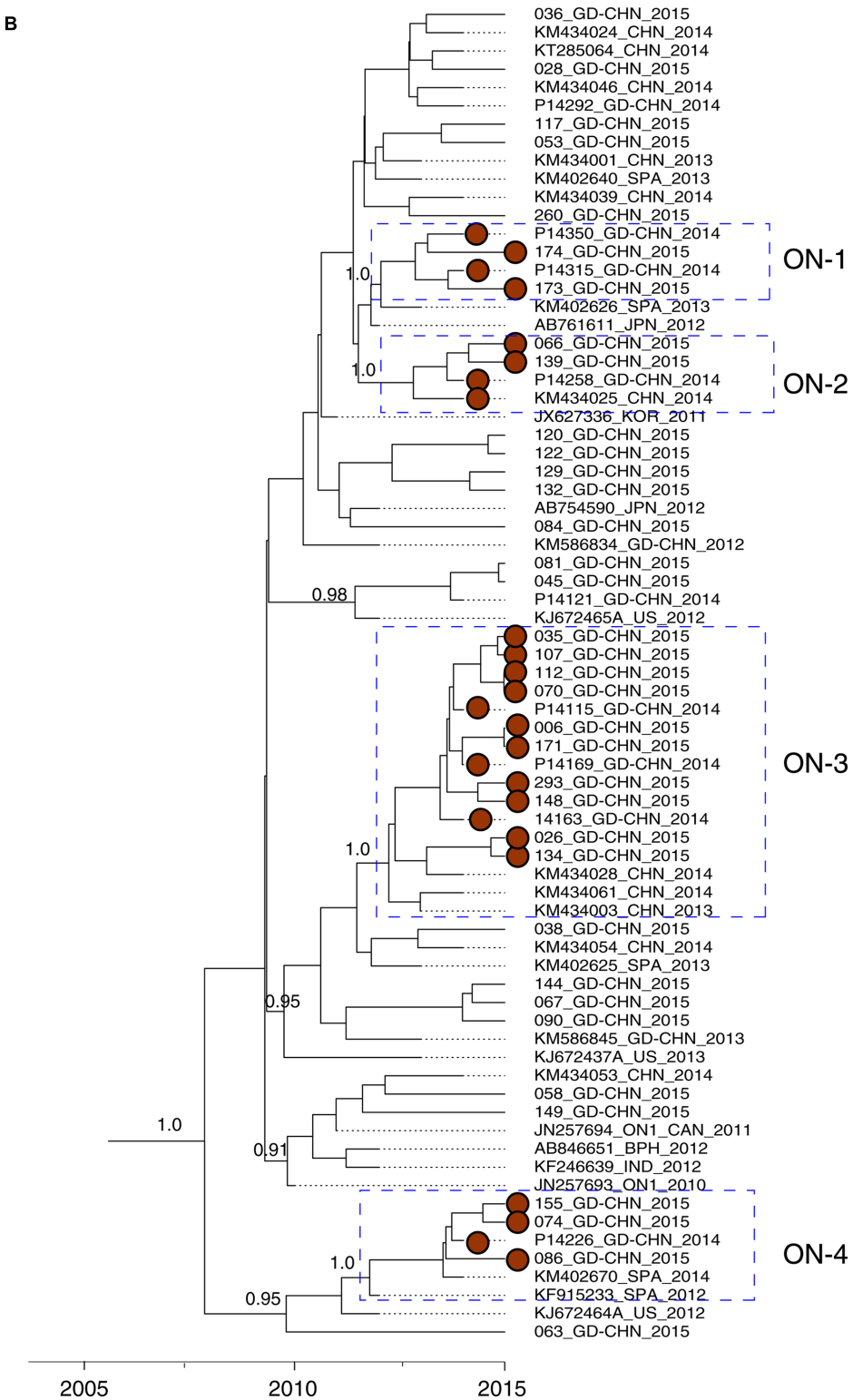


FIGURE 3 | Molecular clock phylogeny of *G* gene sequences from Guangdong between 2008 and 2015. Phylogeny was estimated using maximum Clade Credibility (MCC) on the basis of *G* gene sequences obtained from Guangdong between 2008 and 2015 and closely related sequences from other regions. Posterior values of interested clusters are shown next to nodes. Guangdong RSV-A strains in interested clusters are highlighted with dots filled with colors. **(A)** Viral strains belong to the new emerged ON1 genotype is shown as a triangle. The interested clusters of RSV-A strains with high posterior values are highlighted with boxes and denoted. **(B)** Phylogeny of ON1 genotype RSV-A viruses. The interested clusters are highlighted and denoted.

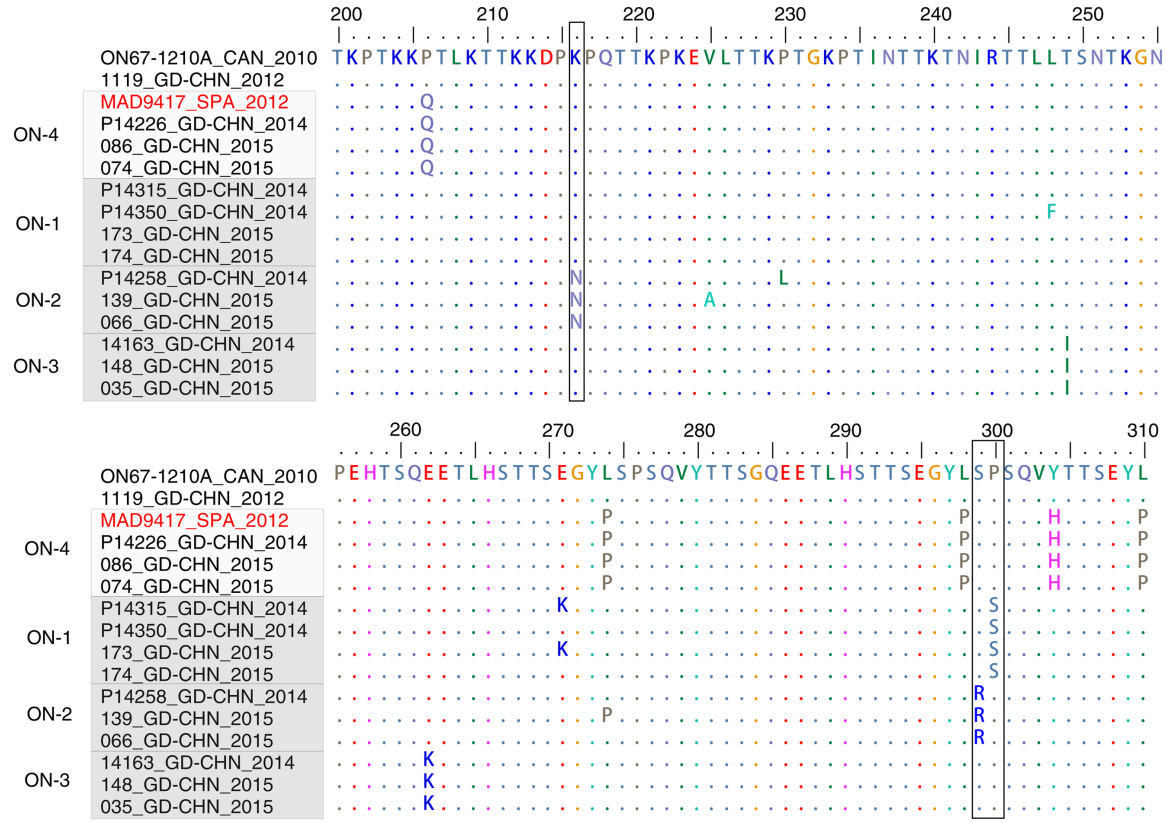


FIGURE 4 | G-protein sequences alignment of ON1 genotype viruses. Representative Guangdong ON1 RSV-A strains and closely related strains are included. The highly variable region in the C-terminal of G-protein sequences (200–310, corresponding to the strain ON67-1210A_2010, accession number JN257693) is shown. The viruses are clustered according to the result of **Figure 3B**. The positions which are exclusively found in Guangdong strains are highlighted in black box.

information on the date and location of sampling were not included. As a consequence, some genotypes like GA3, GA4, and GA6, which may include one or a few sequences, were not represented in the phylogenetic tree. The importance of these sequences is limited because these genotypes (present as minor clades in phylogenies of partial *G* gene sequences) were rarely detected after 2010 in epidemiological studies (Trento et al., 2015).

At the global scale, different circulation patterns are observed for different genotypes of RSV-A. Genotypes such as GA1, GA4, and GA5, are more endemic and display limited epidemic activity throughout their history (**Figure 2**). All appear to have originated and mainly circulated in America, with a few sequences identified on other continents, although this conclusion will not be reliable if early molecular surveillance of RSV was strongly biased toward infections in the USA. However, a distinct pattern is observed for the GA2 genotype, which appears to be more geographically widespread and epidemically active (Eshaghi et al., 2012; Houspie et al., 2013; Trento et al., 2015). Molecular clock analysis reveals that circulation of RSV-A can be classified into three distinct periods (**Figure 2**). Before 1990, RSV-A most sequenced infection cases were mainly identified in America. Between 1990 and 2005, driven by the emergence of GA2, RSV-A viruses were increasingly detected in countries on different continents and temporary co-circulation of different genotypes is observed. The most significant feature of RSV-A molecular epidemiology is the predominance of GA2 genotype after 2005 (Eshaghi et al., 2012; Houspie et al., 2013; Agoti et al., 2014; Liu et al., 2014; Pierangeli et al., 2014; Duvvuri et al., 2015). The prevalence of the GA2 genotype in a population may inhibit the circulation of other genotypes, such as GA7 and GA5, which now are rarely detected, even in America. One possible explanation for this could be the immune cross-protection in a in population generated by GA2 infection. A recent *in vitro* study by Treno et al. (2015) suggested that an antibody (MON-3-88) generated by GA2 virus infection exhibited a broad reactivity to other genotypes including GA3, GA5, and GA7, but not to GA1.

Our study benefits from the continuous surveillance of RSV in Guangdong, enabling us to investigate the genetic diversity of RSV-A across different epidemic seasons in a defined region. Our results indicate that local RSV-A epidemics are caused by a combination of local virus persistence and repeated re-introductions from external locations.

As studies of influenza A viruses have shown (Nelson et al., 2006, 2007; Russell et al., 2008), a simple test of persistence versus seeding is to examine the phylogenetic relationships of strains

sampled between epidemics. If epidemics are mainly caused by virus persistence, new emerged strains would be descended from, and thus more closely related to, strains from previous epidemic in this area. Conversely, if epidemic strains are related to contemporary strains from outside, the epidemics are more likely caused by virus importations.

The phylogenetic analysis highlights both local and global clusters are identified during RSV-A epidemics in Guangdong (**Figures 3A,B**). In addition, the phylogenetic tree mainly based on the Guangdong RSV-A strains (**Figure 2**) is quite similar with the full phylogeny tree of GA2 genotype at global level (**Figure 3**; Supplementary Figure S1). Some Guangdong strains collected in 2008 are near the root of GA2 genotype while the recent sequences collected between 2011 and 2015 in Guangdong fall into the major contemporary cluster of GA2 genotype. One interpretation of this pattern is that the RSV-A circulation in Guangdong is in equilibrium with global RSV-A virus distribution. An alternative but more complex explanation is that viral strains in Guangdong are under parallel evolution after the early dissemination occurred before 2008. The molecular epidemiology of new emerged ON1 suggests the former explanation should be preferred. The ON1 genotype was first detected in Ontario in the winter of 2010/11 (Eshaghi et al., 2012). Thereafter the genotype has been widely spread and prevalent in South Africa (Agoti et al., 2014), Germany (Prifert et al., 2013), Italy (Pierangeli et al., 2014), and Malaysia (Khor et al., 2013). The ON1 RSV-A virus was detected in Guangdong in 2012 and possesses the G protein with the 100% amino acid sequence identity of the Canadian strain suggesting that the novel sub-genotype ON1 virus detected in Guangdong in 2012 is more likely from external seeding. After the possible early introductions, some ON1 strains collected in Guangdong in 2014 and 2015 possess a few amino acid substitutions in G protein are exclusively found in Guangdong and Beijing city (Cui et al., 2015) of China between 2014 and 2015. The prevalence of ON1 genotype in Guangdong also led to the increasing of RSV clinical infection cases and positive rate of RSV in ARIs, 2014–2015. Interestingly, the most recent study from Kenya also suggested both new introductions and local persistence of RSV-A viruses contribute the recurrent of epidemics (Otieno et al., 2016). However, it should be also noted that current phylogenies likely substantially underestimate the contribution of re-introductions to a given location (e.g., Guangdong) due to low levels of sequence sampling from many countries and regions. More representative sampling across the globe, or within a more geographically confined area of interest, will provide more robust results on transmission pattern of RSV-A virus. Another limitation should be emphasized is that our analysis do not have extensive Guangdong RSV-A sequences even though large number of ARIs samples are collected (Supplementary Table S1). This is a common problem for RSV epidemiology studies, partly because a wide range of pathogens can cause ARIs, including bacteria and other viruses besides RSV such as Influenza viruses, parainfluenza virus, adenovirus and human rhinoviruses (van den Hoogen et al., 2001; Azziz-Baumgartner et al., 2012; Kwofie et al., 2012; Marcone et al., 2012). However, the limit surveillance data in Guangdong has already shown some local clusters of RSV with several distinct mutations (**Figures 3** and **4**) and the ON1 prevalence in Guangdong following its emergence in Canada. These results provide the evidence of the joint transmission pattern of RSV-A in Guangdong.

Currently, the RSV GA2 genotype that circulate worldwide has turned into a predominant RSV-A genotype with new variants emerging through a process of continuous evolution that appears somewhat similar to that observed for the global circulation of seasonal influenza viruses (Russell et al., 2008) and human noroviruses (Siebenga et al., 2010; Lu et al., 2016). Our local data suggest RSV-A epidemics in Guangdong are caused by viruses seeded from external regions and viruses persist in local jointly. In this context, for global RSV disease control and vaccine development, much more detailed surveys of RSV-A genetic diversity and evolution in all affected areas, comparable to the surveillance undertaken here for Guangdong province, should be encouraged.

AUTHOR CONTRIBUTIONS

LZ and CK designed the study. LZ, LY, JW, YS, GH, XZ, LL, and HN prepared sample collection and genome sequencing. LY and JL analyzed the data. JL and OP interpreted the data. JL and OP wrote the paper. All authors reviewed the manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China [81501754], Natural Science Foundation of Guangdong Province [2015A030310013], Science and Technology Planning Project of Guangdong Province [2014A020212243], Guangdong Medical Science Foundation [A2016538] and China Scholarship Council [201508440009].

ACKNOWLEDGMENTS

We gratefully acknowledge the authors, originating and submitting laboratories of the sequences from GenBank Database used in the phylogenetic analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01263

REFERENCES

Agoti, C. N., Otieno, J. R., Gitahi, C. W., Cane, P. A., and Nokes, D. J. (2014). Rapid spread and diversification of respiratory syncytial virus genotype ON1, Kenya. *Emerg. Infect. Dis.* 20, 950–959. doi: 10.3201/eid2006.131438

Azziz-Baumgartner, E., Alamgir, A. S., Rahman, M., Homaira, N., Sohel, B. M., Sharker, M. A., et al. (2012). Incidence of influenza-like illness and severe acute respiratory infection during three influenza seasons in Bangladesh, 2008-2010. *Bull. World Health Organ.* 90, 12–19. doi: 10.2471/BLT.11.090209

Bose, M. E., He, J., Shrivastava, S., Nelson, M. I., Bera, J., Halpin, R. A., et al. (2015). Sequencing and analysis of globally obtained human respiratory syncytial virus A and B genomes. *PLoS ONE* 10:e0120098. doi: 10.1371/journal.pone.0120098

Botosso, V. F., Zanotto, P. M., Ueda, M., Arruda, E., Gilio, A. E., Vieira, S. E., et al. (2009). Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog* 5:e1000254. doi: 10.1371/journal.ppat.1000254

Cui, G., Zhu, R., Deng, J., Zhao, L., Sun, Y., Wang, F., et al. (2015). Rapid replacement of prevailing genotype of human respiratory syncytial virus by genotype ON1 in Beijing, 2012-2014. *Infect. Genet. Evol.* 33, 163–168. doi: 10.1016/j.meegid.2015.04.025

Do, L. A., Wilm, A., van Doorn, H. R., Lam, H. M., Sim, S., Sukumaran, R., et al. (2015). Direct whole-genome deep-sequencing of human respiratory syncytial virus A and B from Vietnamese children identifies distinct patterns of inter- and intra-host evolution. *J. Gen. Virol.* 96, 3470–3483. doi: 10.1099/jgv.0.000298

Dong, W., Chen, Q., Hu, Y., He, D., Liu, J., Yan, H., et al. (2016). Epidemiological and clinical characteristics of respiratory viral infections in children in Shanghai, China. *Arch. Virol.* 161, 1907–1913. doi: 10.1007/s00705-016-2866-z

Drummond, A. J., Ho, S. Y., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Duvvuri, V. R., Granados, A., Rosenfeld, P., Bahl, J., Eshaghi, A., and Gubbay, J. B. (2015). Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: global and local transmission dynamics. *Sci. Rep.* 5:14268. doi: 10.1038/srep14268

Edwards, C. J., Suchard, M. A., Lemey, P., Welch, J. J., Barnes, I., Fulton, T. L., et al. (2011). Ancient hybridization and an Irish origin for the modern polar bear matriline. *Curr. Biol.* 21, 1251–1258. doi: 10.1016/j.cub.2011.05.058

Eshaghi, A., Duvvuri, V. R., Lai, R., Nadarajah, J. T., Li, A., Patel, S. N., et al. (2012). Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. *PLoS ONE* 7:e32807. doi: 10.1371/journal.pone.0032807

Fan, R., Fan, C., Zhang, J., Wen, B., Lei, Y., Liu, C., et al. (2016). Respiratory syncytial virus subtype ON1/NA1/BA9 predominates in hospitalized children with lower respiratory tract infections. *J. Med. Virol.* doi: 10.1002/jmv.24619 [Epub ahead of print].

Garcia, O., Martin, M., Dopazo, J., Arbiza, J., Frabasile, S., Russi, J., et al. (1994). Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J. Virol.* 68, 5448–5459.

Glezen, W. P., Taber, L. H., Frank, A. L., and Kasel, J. A. (1986). Risk of primary infection and reinfection with respiratory syncytial virus. *Am. J. Dis. Child.* 140, 543–546.

Gonzalez, P. A., Bueno, S. M., Carreno, L. J., Riedel, C. A., and Kalergis, A. M. (2012). Respiratory syncytial virus infection and immunity. *Rev. Med. Virol.* 22, 230–244. doi: 10.1002/rmv.1704

Henderson, F. W., Collier, A. M., Clyde, W. A. Jr., and Denny, F. W. (1979). Respiratory-syncytial-virus infections, reinfections and immunity. *A* prospective, longitudinal study in young children. *N. Engl. J. Med.* 300, 530–534. doi: 10.1056/NEJM197903083001004

Hirano, E., Kobayashi, M., Tsukagoshi, H., Yoshida, L. M., Kuroda, M., Noda, M., et al. (2014). Molecular evolution of uman respiratory syncytial virus attachment glycoprotein (G) gene of new genotype ON1 and ancestor NA1. *Infect. Genet. Evol.* 28, 183–191. doi: 10.1016/j.meegid.2014.09.030

Houspie, L., Lemey, P., Keyaerts, E., Reijmen, E., Vergote, V., Vankeerberghen, A., et al. (2013). Circulation of HRSV in Belgium: from multiple genotype circulation to prolonged circulation of predominant genotypes. *PLoS ONE* 8:e60416. doi: 10.1371/journal.pone.0060416

Khor, C. S., Sam, I. C., Hooi, P. S., and Chan, Y. F. (2013). Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. *Infect. Genet. Evol.* 14, 357–360. doi: 10.1016/j.meegid.2012.12.017

Kimura, H., Nagasawa, K., Tsukagoshi, H., Matsushima, Y., Fujita, K., Yoshida, L. M., et al. (2016). Molecular evolution of the fusion protein gene in human respiratory syncytial virus subgroup A. *Infect. Genet. Evol.* 43, 398–406. doi: 10.1016/j.meegid.2016.06.019

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., and Frost, S. D. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Mol. Biol. Evol.* 23, 1891–1901. doi: 10.1093/molbev/msl051

Kwofie, T. B., Anane, Y. A., Nkrumah, B., Annan, A., Nguah, S. B., and Owusu, M. (2012). Respiratory viruses in children hospitalized for acute lower respiratory tract infection in Ghana. *Virol. J.* 9:78. doi: 10.1186/1743-422X-9-78

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531

Lemey, P., Rambaut, A., Bedford, T., Faria, N., Bielejec, F., Baele, G., et al. (2014). Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog* 10:e1003932. doi: 10.1371/journal.ppat.1003932

Liu, J., Mu, Y., Dong, W., Yao, F., Wang, L., Yan, H., et al. (2014). Genetic variation of human respiratory syncytial virus among children with fever and respiratory symptoms in Shanghai, China, from 2009 to 2012. *Infect. Genet. Evol.* 27, 131–136. doi: 10.1016/j.meegid.2014.07.011

Lu, J., Fang, L., Zheng, H., Lao, J., Yang, F., Sun, L., et al. (2016). The evolution and transmission of epidemic GII.*17* Noroviruses. *J. Infect. Dis.* 214, 556–564. doi: 10.1093/infdis/jiw208

Marcone, D. N., Ellis, A., Videla, C., Ekstrom, J., Ricarte, C., Carballal, G., et al. (2012). Viral etiology of acute respiratory infections in hospitalized and outpatient children in Buenos Aires, Argentina. *Pediatr. Infect. Dis. J.* 32, e105–e110. doi: 10.1097/INF.0b013e31827cd06f

McLellan, J. S., Ray, W. C., and Peeples, M. E. (2013). Structure and function of respiratory syncytial virus surface glycoproteins. *Curr. Top. Microbiol. Immunol.* 372, 83–104. doi: 10.1007/978-3-642-38919-1_4

Minin, V. N., Bloomquist, E. W., and Suchard, M. A. (2008). Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* 25, 1459–1471. doi: 10.1093/molbev/msn090

Mufson, M. A., Orvell, C., Rafnar, B., and Norrby, E. (1985). Two distinct subtypes of human respiratory syncytial virus. *J. Gen. Virol.* 66(Pt 10), 2111–2124. doi: 10.1099/0022-1317-66-10-2111

Nair, H., Nokes, D. J., Gessner, B. D., Dherani, M., Madhi, S. A., Singleton, R. J., et al. (2010). Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* 375, 1545–1555. doi: 10.1016/S0140-6736(10)60206-1

Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A., and Holmes, E. C. (2007). Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. *PLoS Pathog* 3: e131. doi: 10.1371/journal.ppat.0030131

Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A., Taylor, J., George, K. S., et al. (2006). Stochastic processes are key determinants of short-term evolution in influenza a virus. *PLoS Pathog* 2:e125. doi: 10.1371/journal.ppat.0020125

Otieno, J. R., Agoti, C. N., Gitahi, C. W., Bett, A., Ngama, M., Medley, G. F., et al. (2016). Molecular evolutionary dynamics of respiratory syncytial virus group A in recurrent epidemics in Coastal Kenya. *J. Virol.* 90, 4990–5002. doi: 10.1128/JVI.03105-15

Peret, T. C., Hall, C. B., Schnabel, K. C., Golub, J. A., and Anderson, L. J. (1998). Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J. Gen. Virol.* 79(Pt 9), 2221–2229. doi: 10.1099/0022-1317-79-9-2221

Pierangeli, A., Trotta, D., Scagnolari, C., Ferreri, M. L., Nicolai, A., Midulla, F., et al. (2014). Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. *Euro. Surveill* 19, 20843. doi: 10.2807/1560-7917.ES2014.19.26.20843

Pretorius, M. A., van Niekerk, S., Tempia, S., Moyes, J., Cohen, C., Madhi, S. A., et al. (2013). Replacement and positive evolution of subtype A and B respiratory

syncytial virus G-protein genotypes from 1997-2012 in South Africa. *J. Infect. Dis.* 208(Suppl. 3), S227–S237. doi: 10.1093/infdis/jit477

Prifert, C., Streng, A., Krempl, C. D., Liese, J., and Weissbrich, B. (2013). Novel respiratory syncytial virus a genotype, Germany, 2011-2012. *Emerg. Infect. Dis.* 19, 1029–1030. doi: 10.3201/eid1906.121582

Rambaut, A., Lam, T. T., Max Carvalho, L., and Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2:vew007. doi: 10.1093/ve/vew007

Russell, C. A., Jones, T. C., Barr, I. G., Cox, N. J., Garten, R. J., Gregory, V., et al. (2008). The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320, 340–346. doi: 10.1126/science.1154137

Scott, P. D., Ochola, R., Ngama, M., Okiro, E. A., James Nokes, D., Medley, G. F., et al. (2006). Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya. *J. Infect. Dis.* 193, 59–67. doi: 10.1086/498246

Shapiro, B., Rambaut, A., and Drummond, A. J. (2006). Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* 23, 7–9. doi: 10.1093/molbev/msj021

Siebenga, J. J., Lemey, P., Kosakovsky Pond, S. L., Rambaut, A., Vennema, H., and Koopmans, M. (2010). Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. *PLoS Pathog* 6:e1000884. doi: 10.1371/journal.ppat.1000884

Storey, S. (2010). Respiratory syncytial virus market. *Nat. Rev. Drug Discov.* 9, 15–16. doi: 10.1038/nrd3075

Tan, L., Lemey, P., Houspie, L., Viveen, M. C., Jansen, N. J., van Loon, A. M., et al. (2012). Genetic variability among complete human respiratory syncytial virus subgroup A genomes: bridging molecular evolutionary dynamics and epidemiology. *PLoS ONE* 7:e51439. doi: 10.1371/journal.pone.0051439

Trento, A., Abrego, L., Rodriguez-Fernandez, R., Gonzalez-Sanchez, M. I., Gonzalez-Martinez, F., Delfraro, A., et al. (2015). Conservation of G-Protein epitopes in respiratory Syncytial virus (Group A) despite broad genetic diversity: is antibody selection involved in virus evolution? *J. Virol.* 89, 7776–7785. doi: 10.1128/JVI.00467-15

Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256

van den Hoogen, B. G., de Jong, J. C., Groen, J., Kuiken, T., de Groot, R., Fouchier, R. A., et al. (2001). A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat. Med.* 7, 719–724. doi: 10.1038/89098

Welliver, R. C. (2003). Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. *J. Pediatr.* 143(Suppl. 5), S112–S117. doi: 10.1067/S0022-3476(03)00508-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zou, Yi, Wu, Song, Huang, Zhang, Liang, Ni, Pybus, Ke and Lu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 05 February 2016
doi: 10.3389/fmicb.2016.00085

Genetic Diversity and Positive Selection Analysis of Classical Swine Fever Virus Envelope Protein Gene E2 in East China under C-Strain Vaccination

Dongfang Hu†, Lin Lv†, Jinyuan Gu, Tongyu Chen, Yihong Xiao and Sidang Liu**

Department of Animal Science and Technology, Shandong Agricultural University, Tai'an, China

OPEN ACCESS

Edited by:
Akio Adachi, Tokushima University Graduate School, Japan

Reviewed by:
Stefan Vilcek, University of Veterinary Medicine and Pharmacy in Košice, Slovakia
Myung-Hee Kwon, Ajou University School of Medicine, South Korea

***Correspondence:**
Yihong Xiao
xiaoyihong01@163.com;
Sidang Liu
liusid@sdau.edu.cn

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *10 December 2015*
Accepted: *18 January 2016*
Published: *05 February 2016*

Citation:
Hu D, Lv L, Gu J, Chen T, Xiao Y and Liu S (2016) Genetic Diversity and Positive Selection Analysis of Classical Swine Fever Virus Envelope Protein Gene E2 in East China under C-Strain Vaccination. Front. Microbiol. 7:85. doi: 10.3389/fmicb.2016.00085

Classical swine fever virus (CSFV) causes an economically important and highly contagious disease of pigs worldwide. C-strain vaccination is one of the most effective ways to contain this disease. Since 2014, sporadic CSF outbreaks have been occurring in some C-strain vaccinated provinces of China. To decipher the disease etiology, 25 CSFV E2 genes from 169 clinical samples were cloned and sequenced. Phylogenetic analyses revealed that all 25 isolates belonged to subgenotype 2.1. Twenty-three of the 25 isolates were clustered in a newly defined subgenotype, 2.1d, and shared some consistent molecular characteristics. To determine whether the complete E2 gene was under positive selection pressure, we used a site-by-site analysis to identify specific codons that underwent evolutionary selection, and seven positively selected codons were found. Three positively selected sites (amino acids 17, 34, and 72) were identified in antigenicity-relevant domains B/C of the amino-terminal half of the E2 protein. In addition, another positively selected site (amino acid 200) exhibited a polarity change from hydrophilic to hydrophobic, which may change the antigenicity and virulence of CSFV. The results indicate that the circulating CSFV strains in Shandong province were mostly clustered in subgenotype 2.1d. Moreover, the identification of these positively selected sites could help to reveal molecular determinants of virulence or pathogenesis, and to clarify the driving force of CSFV evolution in East China.

Keywords: classical swine fever virus, genetic diversity, phylogenetic analysis, positive selection, subgenotype 2.1d

INTRODUCTION

Classical swine fever (CSF), previously known as hog cholera, is an economically important, highly contagious disease of pigs that is classified as a notifiable disease by the Office International des Epizooties (Jiang et al., 2013). CSF is characterized by fever and hemorrhage with an acute or chronic course (Luo et al., 2011). CSF was first recognized in Tennessee, USA, in 1810, and then

Abbreviations: AA, amino acid; CSFV, classical swine fever virus; dN, non-synonymous substitution rate; dS, synonymous rate; H&E, hematoxylin and eosin; nt, nucleotide; NTRs, non-translated regions; ORF, open reading frame; PCV2, porcine circovirus type 2; PRRSV, porcine reproductive and respiratory virus; PRV, porcine pseudorabies virus.

rapidly spread throughout the world (Edwards et al., 2000). As a result of systemic immunizations with live attenuated vaccines and/or strict epidemiological surveillance, CSF had been controlled and successfully eradicated from domestic pigs in some countries and regions, such as Australia, New Zealand, North America, and Western Europe (Paton and Greiser-Wilke, 2003; Ji et al., 2015). However, it still significantly affects swine production in Asia, South America, Eastern Europe, and parts of the former Soviet Union (Ji et al., 2015).

The causative agent, CSFV, is a member of the genus *Pestivirus* within the family *Flaviviridae* (Lowings et al., 1996). The positive-sense, single-stranded RNA CSFV genome is 12.3 kb in length, and it comprises one large ORF that is flanked by two NTRs (Rumenapf et al., 1991; Tautz et al., 2015). The ORF codes a 3898-AA polyprotein that is co- and post-translationally processed by cellular and viral proteases into four structural (C, Erns, E1, and E2) and eight non-structural proteins in the order NH2–(Npro-C-Erns-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B)–COOH (Rumenapf et al., 1991; Chang et al., 2010). The E2 protein is the main immunogen of CSFV, and it induces the production of neutralizing antibodies that provide protection against lethal challenge (Beer et al., 2015); it also plays multiple roles in the viral life cycle, and it mediates the entry of the virus into host cells (Sanchez et al., 2008; Shen et al., 2011).

The various isolates of CSFV consist of one serotype, reflecting a narrow range of evolutionary divergence (Vanderhallen et al., 1999; Deng et al., 2005). Therefore, genetic typing of the virus has been used to understand the evolution and spread of viruses, and the origins of disease outbreaks (Deng et al., 2005). 5′-NTR (96 nt), partial E2 (190 nt), and NS5B (409 nt) sequence similarities are extensively used for genetic analyses and to study viral diversity (Lowings et al., 1996; Greiser-Wilke et al., 1998; Paton et al., 2000). Recently, the full-length E2 coding sequence (1,119 nt) was also demonstrated to be reliable in detailed phylogenetic analyses (Postel et al., 2012; Beer et al., 2015). Analyses using these three or four regions have similarly classified CSFV into three genotypes, each with three to four subgenotypes (Lowings et al., 1996; Paton et al., 2000; Deng et al., 2005). Thus far, subgenotypes 1.1–1.4, 2.1–2.3, and 3.1–3.4 can be differentiated (Lowings et al., 1996; Paton et al., 2000; Postel et al., 2013).

Determining the selection pressures that have shaped the genetic variation of viruses is a major part of many molecular evolution studies (Kosakovsky Pond and Frost, 2005). A powerful method for studying adaptive molecular evolution is the use of a codon substitution model to identify AA sites where the dN exceeds the dS in a maximum likelihood context (Anisimova et al., 2001; Shen et al., 2011). Estimates of dN that are significantly different from dS provide convincing evidence for non-neutral evolution (Kosakovsky Pond and Frost, 2005). In viruses, the AAs at the interacting sites between envelope proteins and host molecules are continuously evolving under positive selection (Shen et al., 2011).

Since late 2014 in many regions of Shandong province in East China, a CSF epidemic, which is characterized by abortions and stillbirths of sows, as well as fever, anorexia, skin hemorrhages, and high-mortality among nursery pigs, has been occurring in many pig herds that were immunized with attenuated CSFV vaccines (the C strain, Hog Cholera Lapinized Virus). Most pigs in Shandong are vaccinated according to the following schemes: sows and boars are vaccinated simultaneously three times per year. Piglets are vaccinated first via an intramuscular injection at 21–28 days of age, and they receive a second vaccination at 7–8 weeks of age. Replacement gilts and boars are then vaccinated at 12–16 weeks, followed by a supplementary immunization before estrus (unpublished data). Here, we conducted a molecular epidemiological survey of 25 CSFV isolates and showed that the circulating CSFV strains in Shandong province were mostly clustered in subgenotype 2.1d. The selection pressures that act on the E2 gene of these new isolates and 120 reference strains were further analyzed to obtain insights into the driving forces of CSFV evolution in swine populations under regular vaccination programs.

MATERIALS AND METHODS

Sample Preparation and Virus Isolation

A total of 169 tissue specimens, including the spleen, lymph nodes, tonsils, brain, lungs, and kidneys, were collected from clinically ill nursery pigs from different pig herds of various sizes in Shandong province from December 2013 to June 2015. The tissue samples were collected in accordance with the guidelines of the Shandong Agricultural University Animal Care and Use Committee (SDAUA-2013-001) and dissected for cryopreservation and fixed in 10% neutral formalin for virus detection and histological examination, respectively. Tissue samples were homogenized in Dulbecco's modified Eagle's medium (Gibco, Grand Island, NY, USA), and then the tissue homogenates were centrifuged at 10,000 × *g* (4°C) for 10 min. Then, the suspension was passed through a 0.22-μm filter (EMD Millipore, Billerica, MA, USA) and transferred to PK-15 cell monolayers. Then, the cells were incubated at 37°C in 5% CO_2 for 3–5 days, and the cultures were harvested and stored at –80°C as viral stocks.

Histological Examination and Polymerase Chain Reaction (PCR) Detection

The formalin-fixed samples were processed and embedded in paraffin. Thin sections of the fixed tissues were stained with H&E and examined microscopically. Viral DNA and RNA of the harvested cultures were extracted using the EasyPure viral DNA/RNA kit (TransGen, Beijing, China) according to the manufacturer's instructions for the detection of suspected viruses. Four major pathogens, including CSFV, PRRSV, PRV, and PCV2 were detected by PCR or reverse transcription (RT)-PCR (Hu et al., 2015).

E2 Gene Amplification and Sequencing

Primers based on the published sequence of the CSFV Shimen strain (GenBank accession no. AF092448) were

designed to amplify the complete E2 gene (forward primer: GTAAATATGTGTGTGTTAGACCAGA, reverse primer: GTGT GGGTAATTRAGTTCCCTATCA; Zhang et al., 2015). The viral RNA of CSFV-positive cultures was extracted, and the complete E2 gene was amplified using the EasyScript One-Step RT-PCR SuperMix (TransGen, Beijing, China). Briefly, 6 μL of RNA template, 25 μL of Reaction Mix, 1 μL of Enzyme Mix, and 16 μL of RNase-free water were mixed with 1 μL of each primer (10 μM). One-step RT-PCR was performed using the following conditions: 45°C for 25 min, 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 2 min, followed by a final extension at 72°C for 7 min. PCR/RT-PCR products were analyzed by 1% agarose gel electrophoresis. Target fragments were excised from the gels for purification using the Gel Extraction Kit (Tiangen, Beijing, China). Purified PCR products were cloned into the pMD18-T vector (TaKaRa, Beijing, China). Recombinant clones and the forward and reverse primers were sent to Sangon Bioscience (Shanghai, China) for sequencing.

Phylogenetic Analysis of the E2 Gene

The E2 gene sequences that were amplified from the clinical samples (**Table 1**) were aligned with 120 sequences in GenBank (Supplementary Table S1), and phylogenetic trees were constructed using MEGA 6.0 software[1] by the maximum likelihood method based on the Tamura–Nei model (Tamura and Nei, 1993; Tamura et al., 2013). Bootstrap values were estimated for 1,000 replicates. Trees were determined based on the full-length E2 sequence (1,119 nt) and a partial E2 sequence (190 nt) (Lowings et al., 1996).

Identities and AA Substitution Analysis of the E2 Gene/Protein

The nt and AA sequence identities of the 25 new CSFV isolates and eight representative CSFV isolates, including Shimen (AF092448, 1.1), SXCDK (GQ923951, 2.1a), HEBZ (GU592790, 2.1b), GDPY2008 (HQ697223, 2.1c), SDQS (JQ001834, 2.1d), LAL290 (KC851953, 2.2), Novska (HQ148061, 2.3), and TWN (AY646427, 3.4), were calculated using the MegAlign module (Clustal W method) of the Lasergene package (DNASTAR Inc., Madison, WI, USA). The AA substitutions of the new isolates were compared with those of the representative CSFV isolates, which included three genotypes (1.1–1.4, 2.1–2.3, and 3.4).

Selection Pressure Analysis of the E2 Gene

An analysis of the selection pressure acting on the codons of the E2 envelope protein, including the 25 new isolates and 120 reference strains, was conducted using the HyPhy open-source software package available at the datamonkey web-server[2] (Delport et al., 2010). The level of positive selection was estimated using five different approaches: single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), internal fixed effects likelihood (IFEL), mixed effects model of evolution (MEME), and fast unbiased Bayesian approximation (FUBAR) (Sharma et al., 2013). The best nucleotide substitution model for different datasets, as determined via the available tool on the datamonkey server, was used in the analysis.

[1]http://www.megasoftware.net/

[2]http://www.datamonkey.org/

RESULTS

Gross and Histological Lesions of CSF-Suspected Cases

Systematic necropsies were performed on pigs with clinical signs of CSF, including fever, anorexia, diffuse hemorrhage of the skin (**Figure 1A**), and conjunctivitis. Obvious hemorrhagic spots were found on the surface of the epicardium (**Figure 1B**). Scattered hemorrhagic infarcts were observed on the edge of the spleen (**Figure 1C**). Multiple lymph nodes were hemorrhagic and turgid (**Figure 1D**). The renal cortex were densely covered with petechial hemorrhages (**Figure 1E**). A mixture of small and large hemorrhagic spots, as well as ulcers, was seen on the surface of the gastric mucosa (**Figure 1F**). Histological examination mainly confirmed viral encephalitis, hemorrhages of many tissues, and necrotic foci of lymphoid tissues. The brain tissue exhibited typical viral encephalitis with lymphocyte infiltration around the small blood vessels (**Figure 1G**), as well as the proliferation of glial cells (**Figure 1H**). The histological structure of the spleen was disordered and characterized by necrosis, hemorrhage, and depletion of lymphocytes (**Figure 1I**). The lymph nodes showed hemorrhagic necrotizing lymphadenitis with necrotic lymphocytes and hyperplastic reticular cells (**Figure 1J**). The glomerulus and mesenchyme were hemorrhagic (**Figure 1K**).

Pathogens Detected in the Clinical Samples

The PCR/RT-PCR results showed that 25 of the 169 tissue specimens collected from different herds were positive for CSFV. Among the 25 samples, 12 samples were positive for PCV2, five for PRV, and four for PRRSV (data not shown). All 25 amplified E2 genes were sequenced and submitted to GenBank (**Table 1**).

Phylogenetic Analysis of the E2 Gene

A total of 145 full-length E2 gene (1,119 nt) sequences and 145 corresponding partial E2 gene (190 nt) sequences, including the sequences of the 25 new isolates, were used to construct phylogenetic trees (**Figure 2**). The analysis resulted in a classification of all 145 CSFVs into three main groups (genotypes 1–3) containing eight subgroups (1.1–1.4, 2.1–2.3, and 3.4; **Figure 2**).

Of the 25 new isolates, 21 isolates that were isolated in 2015 (SDHZ-15, SDJNi2-15, SDLY-15, SDTA4-15, SDSK-15, SDXT-15, SDMZ1-15, SDXLS-15, SDZB-15, SDTA3-15, SDMZ2-15, SDJNi4-15, SDLY-15, SDLW2-15, SDWK-15, SDZB2-15, SDJNi5-15, SDTA2-15, SDJNi1-15, SDLW1-15, and SDJNi3-15), two previously isolated strains (SDTA1-13 in 2013 and SDJNa-14 in 2014), and seven previously sequenced isolates [SDQS11 (JQ001834), ZS1-08 (FJ607779), Zj0801 (FJ529205), ZJ7.2005

TABLE 1 | Characteristics of the 25 new isolates and GenBank accession no. of E2 genes (1,119 nt).

No.	Isolates name	Specimen	Time	Accession no.
1	SDJNi1-15	Spleen, lymph nodes	2015.04.25	KT953587
2	SDJNi2-15	Brain, tonsil, lung	2015.03.13	KT953588
3	SDJNa-14	Brain, tonsil, lung	2014.04.12	KT953589
4	SDZB2-15	Spleen, lymph nodes	2015.03.20	KT953590
5	SDTA2-15	Tonsil, lung	2015.04.09	KT953591
6	SDTA1-13	Spleen, lymph nodes, kidneys	2013.12.25	KT953592
7	SDLW1-15	Tonsil, lung	2015.04.01	KT953593
8	SDLY-15	Spleen, tonsil, lung	2015.04.07	KT953594
9	SDTA3-15	Tonsil, lung	2015.04.09	KT953595
10	SDJNi3-15	Spleen, lymph nodes	2015.03.06	KT953596
11	SDLW2-15	Tonsil, lung	2015.03.05	KT953597
12	SDJNi4-15	Tonsil, lung	2015.04.29	KT953598
13	SDMZ1-15	Spleen, lymph nodes	2015.04.25	KT953599
14	SDJNi5-15	Brain, tonsil, lung	2015.04.25	KT953600
15	SDJNi6-15	Tonsil, lung	2015.04.25	KT953601
16	SDWK-15	Tonsil, lung	2015.04.15	KT953602
17	SD19-15	Spleen, lymph nodes	2015.03.11	KT953603
18	SDLY-14	Spleen, lymph nodes, tonsil	2014.12.30	KT953604
19	SDZB-15	Tonsil, brain, kidneys	2015.06.03	KT953605
20	SDMZ2-15	Spleen, lymph nodes	2015.06.09	KT953606
21	SDHZ-15	Tonsil, lung	2015.06.08	KT953607
22	SDSK-15	Spleen, lymph nodes	2015.06.08	KT953608
23	SDXLS-15	Spleen, lymph nodes	2015.06.15	KT953609
24	SDTA4-15	Brain, tonsil, lung	2015.06.10	KT953610
25	SDXT-15	Spleen, lymph nodes, tonsil	2015.06.03	KT953611

(DQ907714), HuZ2-05 (EF683606), SX-04 (EF683623), and SH2-05(EF683621)] belonged to the new subgenotype 2.1d (Zhang et al., 2015). The remaining two new isolates, SD19-15 and SDJNi6-15, were clustered in subgenotype 2.1b. Phylogenetic trees based on the two different gene sequences, including the 145 full-length E2 gene sequences (**Figure 2A**) and 145 partial E2 gene sequences (**Figure 2B**), produced similar results. It is evident that all of the recently isolated CSFV strains in Shandong province were surprisingly divergent from the Shimen reference strain and the vaccine strain HCLV, and that the subgenotype 2.1 CSFV strains (mainly subgenotype 2.1d) predominated in more recent CSF epidemics in Shandong province in East China.

Site Mutation Analysis of the E2 Gene

The E2 gene of the 25 new isolates is 1,119 nt long, encoding a 373-AA protein. When compared with each of the eight reference strains (**Table 2**), the 25 newly isolated strains shared the lowest nt identities (81.7–82.5%) and AA identities (88.5–90.1%) with the TWN strain (subgenotype 3.4). The new isolates shared the highest nt and AA sequence similarities with 2.1 reference strains. When compared with each of the four subgenotypes of genotype 2.1, the two new isolates, SD19-15 and SDJNi6-15, had the highest nt identity (94.2%) and AA identities (96.2 and 97.1%, respectively) with the 2.1b reference strain HEBZ, while the other 23 new isolates shared the highest nt identities (95.9–97.5%) and AA identities (96.2–98.4%) with the 2.1d reference strain SDQS (**Table 2**, Supplementary Tables S2 and S3). In addition, the 25 new isolates had greater similarities to subgenotype 2.1b isolates than to either subgenotype 2.1a or 2.1c isolates, indicating a high similarity between subgenotypes 2.1b and 2.1d; these results are in accordance with the report by Zhang et al. (2015).

Compared with the reference strains, the two 2.1b new isolates, SD19-15 and SDJNi6-15, showed no characteristic AA substitutions, while the other 23 new isolates, which belonged to the 2.1d subgenotype, had some unique characteristics (**Figure 3**). Compared with all of the other isolates, the new 2.1d isolates, as well as four of the 2.1d reference strains (DQ907714, FJ529205, FJ607779, and JQ001834) showed consistent AA substitutions, including an R at position 31 (R^{31}), S^{34}, I^{56}, K^{303}, and A^{331}. The subgenotype 2.1d isolates also showed unique AA substitutions, including G/D/N^{36}S, D^{97}N, K/N^{159}R, and V/M/I^{168}A. In addition, some subgenotype 2.1d isolates had two AA substitutions at positions 200 (Q^{200}L) and 205 (R^{205}K) compared with subgenotype 2.1a, 2.1b, and 2.1c isolates.

Selection Pressure Analysis

A selection pressure analysis of the E2 gene of 145 global CSFV strains revealed seven positively selected sites (AAs 17, 34, 72, 168, 200, 240, and 283) by at least two methods (**Table 3**). The detected positively selected sites were diverse, and most of the sites were hydrophilic sites (**Table 4**). There were no regular changes in polarity of the positively selected AAs, but a change from a polar AA (Q) at position 200 to non-polar AAs (V, P, and L) was observed (**Table 4**).

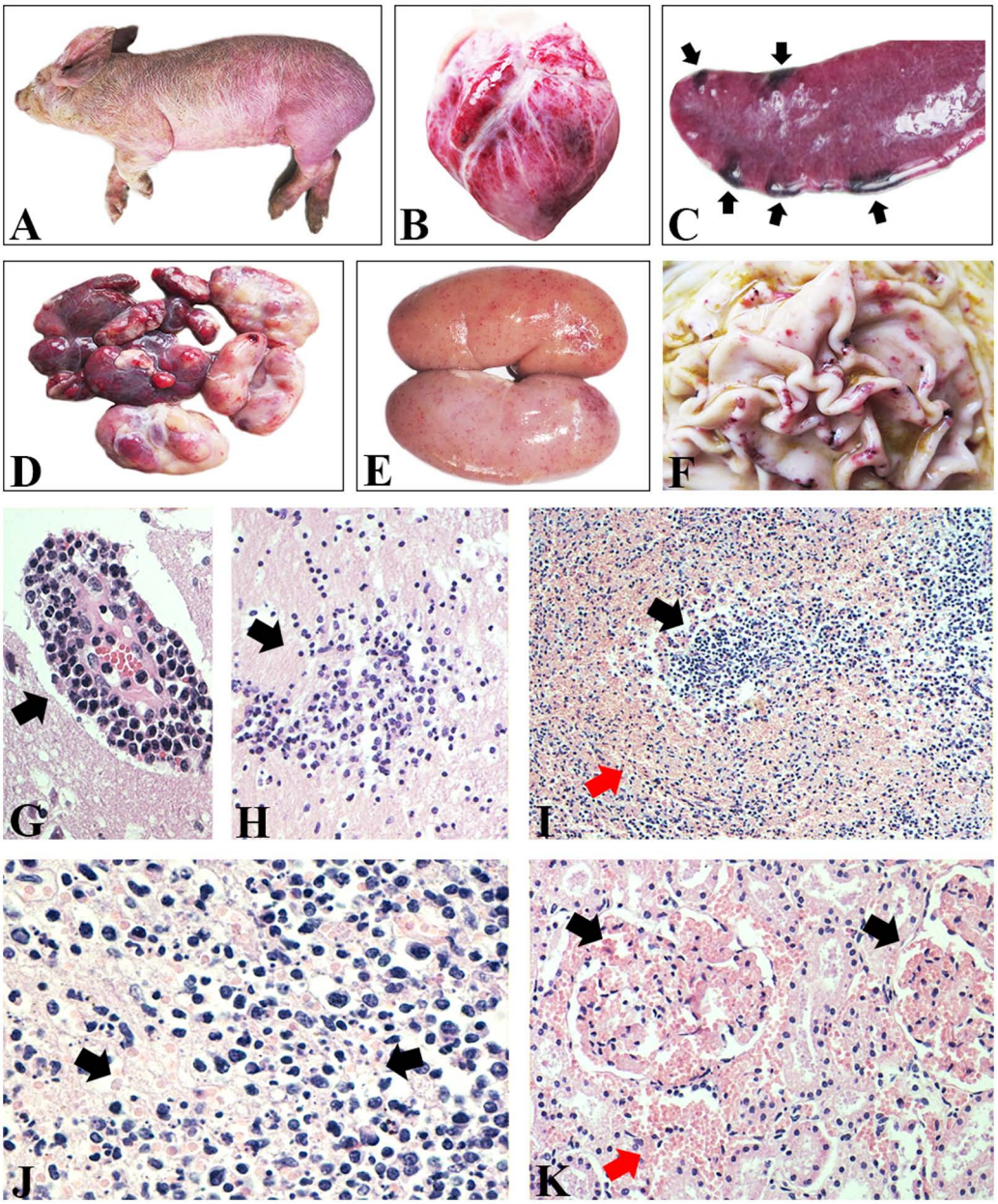


FIGURE 1 | Gross and histological lesions of CSF-suspected pigs. (A) Diffuse hemorrhage of skin. **(B)** Epicardium hemorrhage. **(C)** Infarcts scattered on the edge of spleen. **(D)** Lymph nodes were hemorrhagic and turgid. **(E)** Renal cortex was densely covered with petechial hemorrhages. **(F)** Hemorrhagic spots and ulcer on the surface of gastric mucosa. **(G)** Lymphocyte infiltration around the small blood vessels in brain. H&E stain, ×400. **(H)** Proliferation of glial cells in brain. H&E stain, ×200. **(I)** Histological structure of spleen was disordered and characterized by necrotic lymphocytes and hemorrhage. H&E stain, ×100. **(J)** Lymphoid nodules showed necrotic lymphocytes and hyperplastic reticular cells and hemorrhage. H&E stain, ×400. **(K)** The glomerulus and mesenchyme were hemorrhagic. H&E stain, ×200.

DISCUSSION

In China, a nationwide policy of biannual vaccinations of pigs in the spring and autumn has been performed using the C-strain vaccine, and large-scale outbreaks of CSF have rarely occurred since its introduction (Shen et al., 2011). Some of the cases that occurred were acute, but many cases of CSF were seen as subclinical, causing reproductive failure, neonatal death, or chronic infection in nursery pigs (Luo et al., 2011; Ji et al., 2015). However, in 2014, pigs in some herds in China that were immunized with attenuated CSFV vaccines showed CSF-suspected symptoms (Zhang et al., 2015), and subsequently a similar epidemic unexpectedly occurred in Shandong province, which

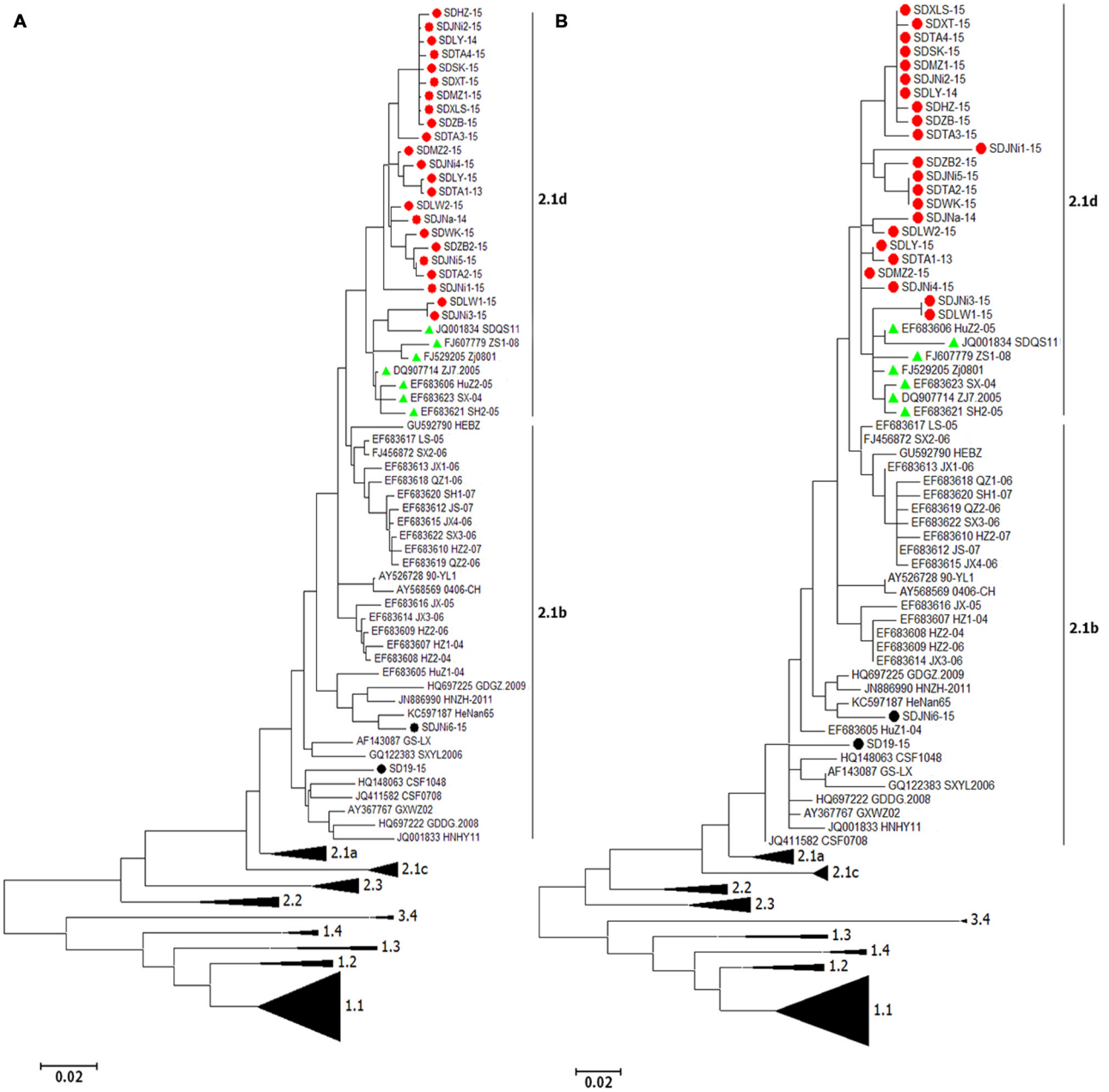


FIGURE 2 | Phylogenetic analysis of different genomic regions of the 25 new isolates and 122 reference strains. (A) Full-length E2 gene sequences. **(B)** Partial E2 gene sequences. 23 newly isolated strains (indicated by ●) accompanied with seven references strains (indicated by ▲) were clustered into subgenotype 2.1d. The other two newly isolated strains (labeled by ●) belonged to subgenotype 2.1b.

caused heavy economic losses. To identify the pathogeneses and pathogens, specimens were collected and systemic examinations were performed, and the CSFV infection status was confirmed.

To further study the molecular epidemiology of CSF, 25 isolated CSFV strains were obtained, and their genetic diversity was analyzed. The full-length E2 gene sequence (1,119 nt), which provides better resolution for phylogenetic analysis than 5′-NTR, partial E2 gene, and NS5B sequences (Blacksell et al., 2004; Sarma et al., 2011; Zhang et al., 2015), was sequenced and examined in this study. Both the full-length E2 sequence and partial E2 sequence showed similar results, as the CSFV isolates could be divided into three genotypes (1, 2, and 3) as well as 11 subgenotypes [1.1–1.4, 2.1 (2.1a, 2.1b, 2.1c, and 2.1d), 2.2, 2.3, and 3.4]. Compared with representative strains of subgenotypes 1.1, 2.1, 2.2, 2.3, and 3.4, the 25 isolates all belonged to subgenotype 2.1, and most of the strains (92%, 23/25) were clustered in the

TABLE 2 | Nucleotide (nt) and AA identities of E2 gene between the 25 new isolates and other eight representative CSFV isolates (%).

Identities	Shimen (1.1)	SXCDK (2.1a)	HEBZ (2.1b)	GDPY2008 (2.1c)	SDQS (2.1d)	LAL290 (2.2)	Novska (2.3)	TWN (3.4)
nt	83.1–84.2	90.5–92.0	94.2–96.3	90.3–92.1	93.7–97.5	85.3–86.7	86.0–87.4	81.7–82.5
AA	88.7–90.9	93.3–96.0	95.4–97.6	94.1–96.8	95.7–98.4	89.5–92.0	90.9–93.6	88.5–90.1

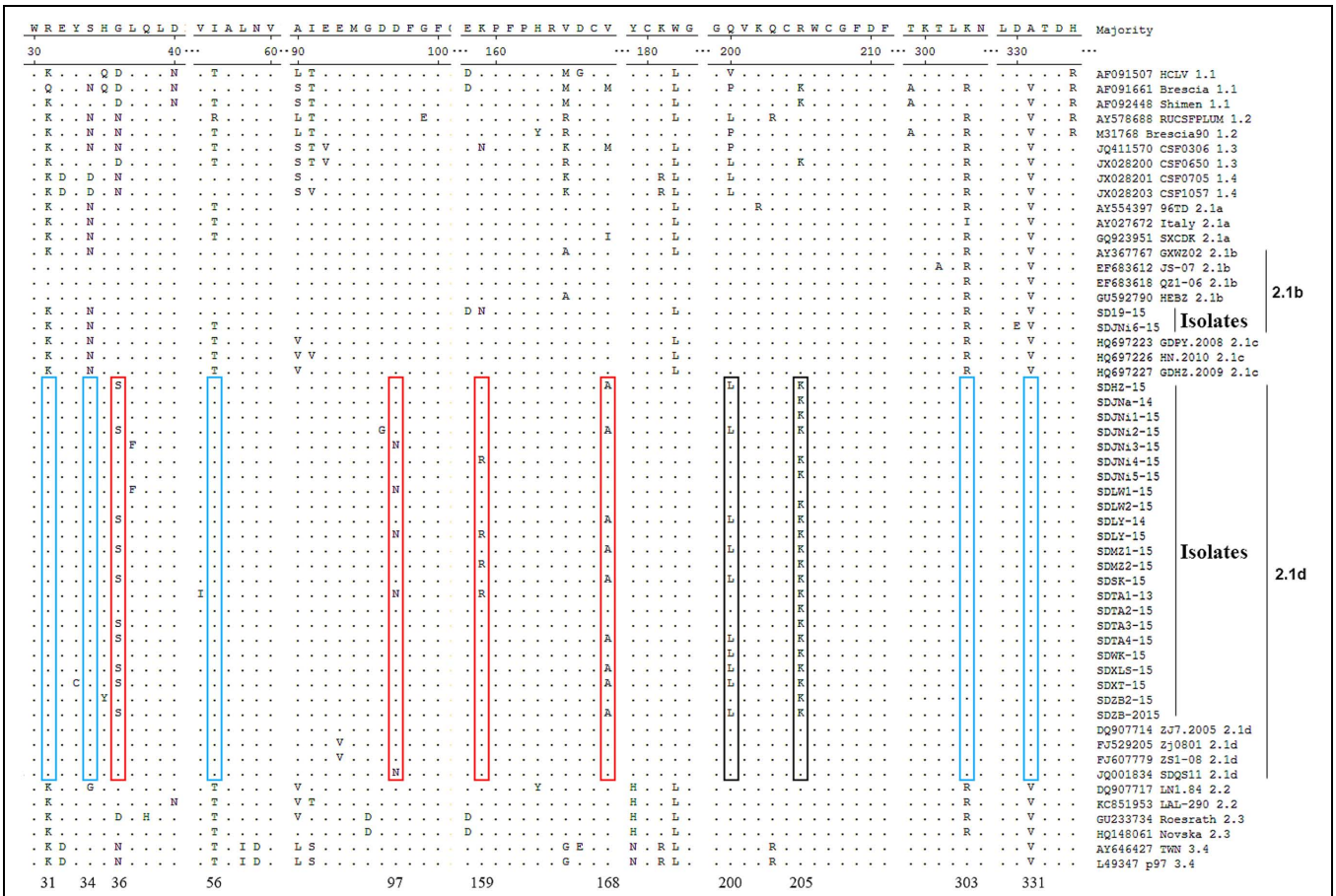


FIGURE 3 | Amino acid (AA) sequences alignments of E2 genes of the 25 new isolates and 29 representative CSFV isolates. Consistent molecular characteristics of 2.1d strains were indicated by blue boxes (□). Unique AA substitutions of 2.1d strains compared with all other strains were indicated by red boxes (□). Unique AA substitutions of 2.1d strains compared with subgenotype 2.1a, 2.1b, and 2.1c strains were indicated by black boxes (□).

newly defined subgenotype 2.1d (**Figure 2**, **Table 2**). High sequence variability is found in mainland China where CSFV subgenotype 1.1, 2.1, 2.2, and 2.3 strains are found, and subgenotype 2.1b has been shown to be the predominant strains within the last 10 years (Tu et al., 2001; Chen et al., 2010a; Beer et al., 2015). In this study, CSF cases caused by a new subgenotype, 2.1d, of CSFV in Shandong province were diagnosed following outbreaks in other provinces (Zhang et al., 2015), and the earliest discovered CSFV isolate, SDTA1-13, which was identified as subgenotype 2.1d in this study, was first isolated in 2013. The results indicate that the new strains may have emerged over a short period of time and spread to several provinces in China, which is worthy of attention because all of the new strains were isolated from CSFV-immunized pigs (Zhang et al., 2015). The pathogenicity, antigenicity, and virulence of the newly defined 2.1d isolates remain unclear, but we speculate that the unique molecular characteristics of the 2.1d isolates may contribute to the adaptive evolution of CSFV under C-strain vaccination, and may be responsible for the unsatisfactory immunoprotection of C-strain vaccinations.

To further study the molecular characteristics of CSFV strains, a selection pressure analysis of E2 AA sequences was performed, and the results showed that the protein mainly underwent purifying selection pressures. RNA viruses are known to have significantly greater mutation rates per site per round of replication than DNA viruses, a difference that is attributed to the error-prone nature of viral RNA-dependent RNA polymerases, and most mutations in coding regions are deleterious (Weiss, 2002; Hughes and Hughes, 2007). A mechanism to decrease

TABLE 3 | Selection pressure analysis of E2 protein (373 codons) of CSFV using SLAC, FEL, IFEL, MEME, and FUBAR methods (www.datamonkey.org).

Protein	Codon*	SLAC		FEL		IFEL		MEME		FUBAR	
		dN-dS	*p*	dN/dS	*p*	dN/dS	*p*	β^+	*p*	β–α	*Post.Pr.*†
E2	17	None	–	None	–	+∞	0.032	None	–	0.118	0.749
	34	4.949	0.149	3.692	0.058	4.492	0.045	1.618	0.076	0.278	0.856
	72	8.142	0.010	+∞	0.004	+∞	0.005	27.412	0.001	0.524	0.971
	168	3.718	0.172	None	–	None	–	4.720	0.035	0.147	0.777
	200	5.768	0.132	None	–	None	–	None	–	0.437	0.820
	240	3.320	0.160	+∞	0.026	+∞	0.072	0.679	0.039	0.190	0.882
	283	2.759	0.198	+∞	0.063	+∞	0.089	218.761	0.020	0.124	0.806

The sites found under positive selection (significance value) by at least two methods are shown.
Codon with $p < 0.2$ level in SLAC method, or with $p < 0.1$ level in FEL method, or with $p < 0.1$ level in IFEL method, or with $p < 0.1$ level in MEME method, or with Post.Pr. $\geq$ 0.7 level in FUBAR method, was considered under positive selection.
**Positively selected sites (AA).*
†*Posterior probability.*

TABLE 4 | Positions and polarities of the positively selected amino acid (AA).

Sites	17	34	72	168	200	240	283
Majority	E^-	S^0	R^+	$V^\times$	Q^0	R^+	R^+
Substitution	K^+, G^0	N^0, D^-, G^0	G^0, E^-, D^-	$M^\times$, $A^\times$, $L^\times$, $I^\times$	$V^\times$, $P^\times$, $L^\times$	G^0, K^+	C^-, H^+, Y^0

The polarities of positively selected AA were indicated at the top-right corner of each AA.
$^\times$*Non-polar AA (hydrophobic).*
$^+$*Positively charged AA (hydrophilic).*
$^-$*Negatively charged AA (hydrophilic).*
^{0}*Uncharged AA (hydrophilic).*

the accumulation of deleterious mutations is essential for RNA viruses to remain stable, and purifying selection provides a useful tool to purge such mutations (Domingo and Holland, 1997). In addition, purifying selection was reportedly more effective in RNA viruses than in DNA viruses (Hughes and Hughes, 2007). Seven positively selected sites were observed in the E2 protein, which is the main immunogen of CSFV. E2 is a type I transmembrane protein with a transmembrane domain in its carboxyl-terminus that is anchored in the viral envelope (Li et al., 2013). The amino-terminal half of the E2 protein, which is an extracellular motif that contains four antigenic domains (A, B, C, and D), was more variable than the carboxyl-terminal half (van Rijn et al., 1994). E2 has a unique architecture consisting of two immunoglobulin-like domains (I and II). Domains D/A map to domain II (AAs 91–168) in the E2 crystal structure, and domains B/C correspond to domain I (AAs 1–90) (Li et al., 2013). Among the detected seven positively selected sites, AAs 17, 34, and 72 belonged to domains B/C. AA 168 belonged to domains D/A. Moreover, the other three sites (AAs 200, 240, and 283) are located in the carboxyl-terminal half of the E2 protein. Domains B/C, which form an independent antigenic unit, are responsible for antigenic specificity among various CSFVs, and the D/A domains of various CSFVs are relatively conserved (van Rijn et al., 1994; Chang et al., 2010). It has been reported that single mutations in the E2 B/C domains could lead to variations in viral neutralization (Chen et al., 2010b). The three positively selected sites found in domains B/C of the amino-terminal half of the E2 protein, which mediates viral entry into target cells, suggest that these changes could be associated with viral escape from neutralizing antibodies, and they could explain the lower severity of the clinical signs that developed in most of the affected animals. The positively selected AA 200 is reportedly necessary for the attenuation of the highly virulent Brescia strain, but the mechanisms mediating this attenuation remain unknown (Risatti et al., 2007; Tang et al., 2008). In this study, we observed a polarity change of AA 200 from hydrophilic to hydrophobic, which may contribute to a change of the antigenicity and virulence of CSFV. The other three positively selected sites (AAs 168, 240, and 283) found in this study are the first to be reported, and their biological significance needs to be further characterized. Understanding the functional importance of these positively selected AAs could help to predict possible changes in virulence, which will aid the study of the mechanism of immune evasion, and prevent CSF in the future.

CONCLUSION

The 25 CSFV isolates from East China were clustered in subgroup 2.1, and most of the isolates, together with some previously sequenced strains, formed the newly defined subgenotype 2.1d, indicating that 2.1d CSFV strains may be predominant epidemic strains in Shandong province. The selection pressure analysis revealed that the envelope protein-encoding E2 gene had

undergone positive selection, and several positively selected sites were identified, which could help to identify the molecular determinants of virulence or pathogenesis, and to clarify the driving force of CSFV evolution in East China. Empirical studies are required to assess the antigenicity and virulence of the 2.1d CSFV strains, as well as the influence of the positively selected AAs identified in this study on CFSV virulence or pathogenesis.

AUTHOR CONTRIBUTIONS

DH and SL contributed to conception and design of the study. YX contributed to design of the study. LL contributed to acquisition and analysis of data. JG and TC contributed to acquisition of data. DH and LL drafted the manuscript. YX and SL critically revised the manuscript.

FUNDING

This research was partially supported by the Open Fund of the State Key Laboratory of Veterinary Etiological Biology (SKLVEB2015KFKT0015).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.00085

REFERENCES

Anisimova, M., Bielawski, J. P., and Yang, Z. (2001). Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol. Biol. Evol.* 18, 1585–1592. doi: 10.1093/oxfordjournals.molbev.a003945

Beer, M., Goller, K. V., Staubach, C., and Blome, S. (2015). Genetic variability and distribution of Classical swine fever virus. *Anim. Health Res. Rev.* 16, 33–39. doi: 10.1017/S1466252315000109

Blacksell, S. D., Khounsy, S., Boyle, D. B., Greiser-Wilke, I., Gleeson, L. J., Westbury, H. A., et al. (2004). Phylogenetic analysis of the E2 gene of classical swine fever viruses from Lao PDR. *Virus Res.* 104, 87–92. doi: 10.1016/j.virusres.2004.02.041

Chang, C. Y., Huang, C. C., Lin, Y. J., Deng, M. C., Chen, H. C., Tsai, C. H., et al. (2010). Antigenic domains analysis of classical swine fever virus E2 glycoprotein by mutagenesis and conformation-dependent monoclonal antibodies. *Virus Res.* 149, 183–189. doi: 10.1016/j.virusres.2010.01.016

Chen, N., Li, D., Yuan, X., Li, X., Hu, H., Zhu, B., et al. (2010a). Genetic characterization of E2 gene of classical swine fever virus by restriction fragment length polymorphism and phylogenetic analysis. *Virus Genes* 40, 389–396. doi: 10.1007/s11262-010-0465-8

Chen, N., Tong, C., Li, D., Wan, J., Yuan, X., Li, X., et al. (2010b). Antigenic analysis of classical swine fever virus E2 glycoprotein using pig antibodies identifies residues contributing to antigenic variation of the vaccine C-strain and group 2 strains circulating in China. *Virol. J.* 7, 378. doi: 10.1186/1743-422X-7-378

Delport, W., Poon, A. F., Frost, S. D., and Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26, 2455–2457. doi: 10.1093/bioinformatics/btq429

Deng, M. C., Huang, C. C., Huang, T. S., Chang, C. Y., Lin, Y. J., Chien, M. S., et al. (2005). Phylogenetic analysis of classical swine fever virus isolated from Taiwan. *Vet. Microbiol.* 106, 187–193. doi: 10.1016/j.vetmic.2004.12.014

Domingo, E., and Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51, 151–178. doi: 10.1146/annurev.micro.51.1.151

Edwards, S., Fukusho, A., Lefevre, P. C., Lipowski, A., Pejsak, Z., Roehe, P., et al. (2000). Classical swine fever: the global situation. *Vet. Microbiol.* 73, 103–119. doi: 10.1016/S0378-1135(00)00138-3

Greiser-Wilke, I., Depner, K., Fritzemeier, J., Haas, L., and Moennig, V. (1998). Application of a computer program for genetic typing of classical swine fever virus isolates from Germany. *J. Virol. Methods* 75, 141–150. doi: 10.1016/S0166-0934(98)00109-8

Hu, D., Zhang, Z., Lv, L., Xiao, Y., Qu, Y., Ma, H., et al. (2015). Outbreak of variant pseudorabies virus in Bartha-K61–vaccinated piglets in central Shandong Province, China. *J. Vet. Diagn. Invest.* 27, 600–605. doi: 10.1177/1040638715593599

Hughes, A. L., and Hughes, M. A. (2007). More effective purifying selection on RNA viruses than in DNA viruses. *Gene* 404, 117–125. doi: 10.1016/j.gene.2007.09.013

Ji, W., Guo, Z., Ding, N. Z., and He, C. Q. (2015). Studying classical swine fever virus: making the best of a bad virus. *Virus Res.* 197, 35–47. doi: 10.1016/j.virusres.2014.12.006

Jiang, D. L., Gong, W. J., Li, R. C., Liu, G. H., Hu, Y. F., Ge, M., et al. (2013). Phylogenetic analysis using E2 gene of classical swine fever virus reveals a new subgenotype in China. *Infect. Genet. Evol.* 17, 231–238. doi: 10.1016/j.meegid.2013.04.004

Kosakovsky Pond, S. L., and Frost, S. D. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* 22, 1208–1222. doi: 10.1093/molbev/msi105

Li, Y., Wang, J., Kanai, R., and Modis, Y. (2013). Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6805–6810. doi: 10.1073/pnas.1300524110

Lowings, P., Ibata, G., Needham, J., and Paton, D. (1996). Classical swine fever virus diversity and evolution. *J. Gen. Virol.* 77(Pt 6), 1311–1321. doi: 10.1099/0022-1317-77-6-1311

Luo, T. R., Liao, S. H., Wu, X. S., Feng, L., Yuan, Z. X., Li, H., et al. (2011). Phylogenetic analysis of the E2 gene of classical swine fever virus from the Guangxi Province of Southern China. *Virus Genes* 42, 347–354. doi: 10.1007/s11262-011-0578-8

Paton, D. J., and Greiser-Wilke, I. (2003). Classical swine fever–an update. *Res. Vet. Sci.* 75, 169–178. doi: 10.1016/S0034-5288(03)00076-6

Paton, D. J., Mcgoldrick, A., Greiser-Wilke, I., Parchariyanon, S., Song, J. Y., Liou, P. P., et al. (2000). Genetic typing of classical swine fever virus. *Vet. Microbiol.* 73, 137–157. doi: 10.1016/S0378-1135(00)00141-3

Postel, A., Schmeiser, S., Bernau, J., Meindl-Boehmer, A., Pridotkas, G., Dirbakova, Z., et al. (2012). Improved strategy for phylogenetic analysis of classical swine fever virus based on full-length E2 encoding sequences. *Vet. Res.* 43, 50. doi: 10.1186/1297-9716-43-50

Postel, A., Schmeiser, S., Perera, C. L., Rodriguez, L. J., Frias-Lepoureau, M. T., and Becher, P. (2013). Classical swine fever virus isolates from Cuba form a new subgenotype 1.4. *Vet. Microbiol.* 161, 334–338. doi: 10.1016/j.vetmic.2012.07.045

Risatti, G. R., Holinka, L. G., Fernandez Sainz, I., Carrillo, C., Lu, Z., and Borca, M. V. (2007). N-linked glycosylation status of classical swine fever virus strain Brescia E2 glycoprotein influences virulence in swine. *J. Virol.* 81, 924–933. doi: 10.1128/JVI.01824-06

Rumenapf, T., Meyers, G., Stark, R., and Thiel, H. J. (1991). Molecular characterization of hog cholera virus. *Arch. Virol. Suppl.* 3, 7–18. doi: 10.1007/978-3-7091-9153-8_2

Sanchez, O., Barrera, M., Rodriguez, M. P., Frias, M. T., Figueroa, N. E., Naranjo, P., et al. (2008). Classical swine fever virus E2 glycoprotein antigen produced in adenovirally transduced PK-15 cells confers complete protection in pigs upon viral challenge. *Vaccine* 26, 988–997. doi: 10.1016/j.vaccine.2007.11.014

Sarma, D. K., Mishra, N., Vilcek, S., Rajukumar, K., Behera, S. P., Nema, R. K., et al. (2011). Phylogenetic analysis of recent classical swine fever virus (CSFV)

isolates from Assam, India. *Comp. Immunol. Microbiol. Infect. Dis.* 34, 11–15. doi: 10.1016/j.cimid.2009.09.005

Sharma, S., Joshi, G., Dash, P. K., Thomas, M., Athmaram, T. N., Kumar, J. S., et al. (2013). Molecular epidemiology and complete genome characterization of H1N1pdm virus from India. *PLoS ONE* 8:e56364. doi: 10.1371/journal.pone.0056364

Shen, H., Pei, J., Bai, J., Zhao, M., Ju, C., Yi, L., et al. (2011). Genetic diversity and positive selection analysis of classical swine fever virus isolates in south China. *Virus Genes* 43, 234–242. doi: 10.1007/s11262-011-0625-5

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tang, F., Pan, Z., and Zhang, C. (2008). The selection pressure analysis of classical swine fever virus envelope protein genes Erns and E2. *Virus Res.* 131, 132–135. doi: 10.1016/j.virusres.2007.08.015

Tautz, N., Tews, B. A., and Meyers, G. (2015). The molecular biology of pestiviruses. *Adv. Virus Res.* 93, 47–160. doi: 10.1016/bs.aivir.2015.03.002

Tu, C., Lu, Z., Li, H., Yu, X., Liu, X., Li, Y., et al. (2001). Phylogenetic comparison of classical swine fever virus in China. *Virus Res.* 81, 29–37. doi: 10.1016/S0168-1702(01)00366-5

van Rijn, P. A., Miedema, G. K., Wensvoort, G., Van Gennip, H. G., and Moormann, R. J. (1994). Antigenic structure of envelope glycoprotein E1 of hog cholera virus. *J. Virol.* 68, 3934–3942.

Vanderhallen, H., Mittelholzer, C., Hofmann, M. A., and Koenen, F. (1999). Classical swine fever virus is genetically stable in vitro and in vivo. *Arch. Virol.* 144, 1669–1677. doi: 10.1007/s007050050622

Weiss, R. A. (2002). Virulence and pathogenesis. *Trends Microbiol.* 10, 314–317. doi: 10.1016/S0966-842X(02)02391-0

Zhang, H., Leng, C., Feng, L., Zhai, H., Chen, J., Liu, C., et al. (2015). A new subgenotype 2.1d isolates of classical swine fever virus in China, 2014. *Infect. Genet. Evol.* 34, 94–105. doi: 10.1016/j.meegid.2015.05.031

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Hu, Lv, Gu, Chen, Xiao and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 April 2017
doi: 10.3389/fmicb.2017.00576

GRIM-19 Restricts HCV Replication by Attenuating Intracellular Lipid Accumulation

*Jung-Hee Kim[1], Pil S. Sung[2†], Eun B. Lee[1], Wonhee Hur[1], Dong J. Park[1], Eui-Cheol Shin[2], Marc P. Windisch[3] and Seung K. Yoon[1]**

[1] The Catholic University Liver Research Center and WHO Collaborating Center of Viral Hepatitis, The Catholic University of Korea, Seoul, South Korea, [2] Laboratory of Immunology and Infectious Diseases, Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea, [3] Hepatitis Research Laboratory, Discovery Biology Department, Institut Pasteur Korea, Seongnam-si, Gyeonggi-do, South Korea

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Koichi Watashi,
National Institute of Infectious Diseases, Japan
Takanobu Kato,
National Institute of Infectious Diseases, Japan
Lynn B. Dustin,
University of Oxford, UK
Stacy M. Horner,
Duke University Medical Center, USA

***Correspondence:**
Seung K. Yoon
yoonsk@catholic.ac.kr

†Present address:
Pil S. Sung,
The Catholic University Liver Research Center and WHO Collaborating Center of Viral Hepatitis, The Catholic University of Korea, Seoul, South Korea

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *18 November 2016*
Accepted: *20 March 2017*
Published: *11 April 2017*

Citation:
Kim J-H, Sung PS, Lee EB, Hur W, Park DJ, Shin E-C, Windisch MP and Yoon SK (2017) GRIM-19 Restricts HCV Replication by Attenuating Intracellular Lipid Accumulation. Front. Microbiol. 8:576. doi: 10.3389/fmicb.2017.00576

Gene-associated with retinoid-interferon-induced mortality 19 (GRIM-19) targets multiple signaling pathways involved in cell death and growth. However, the role of GRIM-19 in the pathogenesis of hepatitis virus infections remains unexplored. Here, we investigated the restrictive effects of GRIM-19 on the replication of hepatitis C virus (HCV). We found that GRIM-19 protein levels were reduced in HCV-infected Huh7 cells and Huh7 cells harboring HCV replicons. Moreover, ectopically expressed GRIM-19 caused a reduction in both intracellular viral RNA levels and secreted viruses in HCVcc-infected cell cultures. The restrictive effect on HCV replication was restored by treatment with siRNA against GRIM-19. Interestingly, GRIM-19 overexpression did not alter the level of phosphorylated STAT3 or its subcellular distribution. Strikingly, forced expression of GRIM-19 attenuated an increase in intracellular lipid droplets after oleic acid (OA) treatment or HCVcc infection. GRIM-19 overexpression abrogated fatty acid-induced upregulation of sterol regulatory element-binding transcription factor-1 (SREBP-1c), resulting in attenuated expression of its target genes such as fatty acid synthase (FAS) and acetyl CoA carboxylase (ACC). Treatment with OA or overexpression of SREBP-1c in GRIM-19-expressing, HCVcc-infected cells restored HCV replication. Our results suggest that GRIM-19 interferes with HCV replication by attenuating intracellular lipid accumulation and therefore is an anti-viral host factor that could be a promising target for HCV treatment.

Keywords: hepatitis C virus, anti-viral host factor, viral replication, lipogenesis, intracellular lipid accumulation

INTRODUCTION

After entry into hepatocytes, hepatitis C virus (HCV) becomes uncoated, and the viral genome is translated into a single polyprotein that is co- and post-translationally processed into structural and non-structural proteins (Gale and Foy, 2005; Lindenbach and Rice, 2005). The HCV non-structural proteins, such as NS3 helicase, NS5A, and NS5B RNA-dependent RNA polymerase, assemble as a replicase complex (RC) that is associated with lipid-rich membrane structures (Aizaki et al., 2004; Lindenbach and Rice, 2005). The newly synthesized viral genomes are packaged into viral particles by the structural proteins, including core, E1, and E2. It has been reported that lipid droplets (LDs) are important organelles for the viral packing step in HCV production (Miyanari et al., 2007).

Moreover, newly assembled HCV particles are observed in close proximity to LDs, indicating that some steps of virus assembly occur near LDs (Miyanari et al., 2007). The resulting virus is released from the hepatocyte in association with host lipoproteins, and therefore, in the blood, HCV is present as a lipoprotein-coated virus (Meredith et al., 2012). Host lipid architectures and molecules involved in lipid metabolism are closely associated with the HCV lifecycle (Suzuki, 2012). Many studies have shown that lots of host factors participate in HCV infection and play important roles in efficient viral replication and propagation (Li et al., 2009). One such host factor is signal transducer and activator of transcription 3 (STAT3) (McCartney et al., 2013; Kong et al., 2016; Vallianou et al., 2016). It was reported that HCV core interacts with and activates STAT3. The interaction induces expression of STAT3-dependent genes, such as Bcl-XL and cyclin-D1, resulting in cellular transformation (Yoshida et al., 2002). Another study has shown that STAT3 enhances HCV replication through positive regulation of microtubule dynamics (McCartney et al., 2013). More recently, it has been demonstrated that HCV NS4B induces the production of reactive oxygen species (ROS) via the endoplasmic reticulum overload response (EOR)-mediated cancer-related STAT3 pathway (Kong et al., 2016). Furthermore, the roles of cellular regulators of STAT3 such as protein inhibitor of activated STAT (PIAS) and suppressor of cytokine signaling (SOCS3) have been investigated in the context of HCV pathogenesis (El-Saadany et al., 2013; Li Q. et al., 2014; Xu et al., 2014; Aslam et al., 2016; Zhao et al., 2016). However, the function of gene-associated with retinoid-interferon-induced mortality 19 (GRIM-19), which is another cellular inhibitor of STAT3, remains largely unexplored in HCV infection.

GRIM-19 was identified as an interferon (IFN)-β- and retinoic acid (RA)-induced gene with pro-apoptotic properties in breast cancer cell lines (Angell et al., 2000). Studies have demonstrated that GRIM-19 targets multiple signaling pathways and plays a critical role in controlling cell death and growth. Overexpression of GRIM-19 induces cell death, and its suppression or inactivation promotes cell growth (Moreira et al., 2011). Regarding the role of GRIM-19 in cancer development, GRIM-19 expression was severely downregulated in a number of primary renal cell carcinomas (Alchanati et al., 2006), as well as in hepatocellular carcinoma (Liu et al., 2014) and oral squamous cell carcinoma (Li M. et al., 2014). Accordingly, upregulation of GRIM-19 can suppress the growth of specific cancers (Li M. et al., 2014; Liu et al., 2014). These tumor-suppressive activities of GRIM-19 may be attributed to its inhibitory role in the function of STAT3. GRIM-19 was shown to suppress STAT3-induced gene expression via direct interaction with the trans-activation domain (TAD) of STAT3 (Nallar et al., 2008). In this way, binding of GRIM-19 to STAT3 induces changes in the intracellular distribution of STAT3 and renders cells sensitive to cell death (Shulga and Pastorino, 2012).

Interestingly, the function of GRIM-19 was reported to be impeded by viral factors of oncogenic viruses (Kalvakolanu et al., 2010). The viral interferon regulatory factors (vIRFs) from human herpesvirus-8 (HHV-8), implicated in cellular transformation, bind to GRIM-19 and block its ability to induce apoptosis (Seo et al., 2002). Similarly, a non-coding 2.7-kb viral RNA (β2.7) produced by human cytomegalovirus (CMV) enters mitochondria and locks GRIM-19 into Complex-I, rendering it incapable of triggering apoptosis (Reeves et al., 2007).

In this study, we investigated the role of GRIM-19 as a host factor restricting HCV infection. We observed that HCV infection downregulates GRIM-19 at the post-transcriptional level and that GRIM-19 overexpression interferes with HCV replication. Regarding the mechanism for these effects, we found that GRIM-19 decreases intracellular lipid accumulation by regulating the expression of the sterol regulatory element-binding transcription factor-1 (SREBP-1c) gene and its downstream genes. These results suggest that GRIM-19 may be an anti-viral host factor that could be exploited for the development of novel antiviral agents.

MATERIALS AND METHODS

Antibodies and Reagents

Mouse monoclonal anti-GRIM-19 antibody was purchased from Abcam (Cambridge, MA, USA). Mouse monoclonal anti-β-actin and mouse monoclonal anti-flag antibodies were obtained from Sigma–Aldrich (St. Louis, MO, USA). Mouse monoclonal anti-HCV core antibody was purchased from Thermo Scientific (Rockford, IL, USA). Mouse monoclonal anti-HCV NS5A antibody was obtained from Virogen (Watertown, MA, USA). Polyclonal antibodies specific to phospho-STAT3, STAT3 (Ser-705), and acetyl CoA carboxylase (ACC) were purchased from Cell Signaling Technology, Inc. (Danvers, MA, USA). Horseradish peroxidase (HRP)-conjugated anti-mouse, anti-rabbit immunoglobulin G (IgG), mouse monoclonal anti-SREBP-1c, mouse monoclonal anti-fatty acid synthase (FAS), and goat polyclonal anti-stearoyl CoA desaturase (SCD) antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Scrambled siRNA and siRNA targeting GRIM-19 were obtained from Santa Cruz Biotechnology.

Clinical Materials and Ethics Statement

Four liver tissues from patients with chronic HCV infection were obtained during surgical procedures such as cholecystectomy, adrenalectomy, or partial liver resection for intrahepatic duct stones (Seoul St. Mary's Hospital, Seoul, South Korea). Three out of four patients had cirrhotic liver, and the other patient had chronic hepatitis without cirrhosis. None of them had history of anti-viral treatment. In addition, four liver tissues without viral hepatitis were also obtained during surgical procedures, and they were described in the previous report (Sung et al., 2015). The study conformed to the current ethical principles of the Declaration of Helsinki and was approved by the Institutional Review Board of both Seoul St. Mary's Hospital and Daejeon St. Mary's Hospital at the Catholic University of Korea. All patients who provided their tissues completed written informed consents before inclusion in the study. Additionally, their personal identifying information was restricted for analysis purposes and is not available to the public.

Cell Culture

Huh7 cells were kindly provided by Dr. Jane C. Moores (The Regent of the University of California, Oakland, CA, USA). Dr. Francis Chisari (The Scripps Research Institute, CA) generously provided Huh7.5.1 cells. The cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS), 1% antibiotics (100 μg/mL of penicillin, 0.25 μg/mL of streptomycin), and 10 μM HEPES in a humidified incubator at 37°C in 5% CO_2.

HCVcc Preparation and Infection

Full-length, infectious HCV RNA of the genotype 2a HCV clone JFH1 was prepared by *in vitro* transcription using a MEGAscript T7 kit (Ambion) and electroporated into Huh7 cells to obtain cell culture-derived HCV (HCVcc) as previously described (Wakita et al., 2005). Huh7 cells were infected with HCVcc at a multiplicity of infection (MOI) of 0.3 by adsorption for 6 h with periodic rocking and then maintained in complete DMEM as previously described (Sun et al., 2012).

HCV Replicon Systems

An HCV subgenomic replicon (SGR) construct (pSGR-JFH1) and an HCV full-genomic replicon (FGR) construct (pFGR-JFH1) were kindly provided by Dr. Takaji Wakita (National Institute of Infectious Diseases, Tokyo, Japan). The constructs were linearized and then used for *in vitro* transcription as described above. Huh7 cell-derived cell lines containing the HCV SGR or HCV FGR were established by transfection of *in vitro*-transcribed HCV subgenomic or HCV full-genomic RNA, followed by selection with 500 μg/mL G418 sulfate as previously described (Date et al., 2007). The selected cell lines were maintained in complete DMEM containing 500 μg/mL G418 sulfate. HCV genotype-3 replicon cells derived from Huh7.5.1 cells were kindly provided from Dr. Sung Key Jang (Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea).

Western Blot Analysis

Huh7 cells and Huh7 cells in which HCV replication occurs were lysed with PRO-PREP Protein Extraction Solution (iNtRon BIOTECHNOLGY) containing protease inhibitors. Total protein content was determined using a Bradford protein assay kit (Bio-Rad Laboratories, Hercules, CA, USA). Thirty micrograms of the extracted proteins were subjected to western blot analysis. The analysis was performed as previously described (Choi et al., 2015). The density of each band was analyzed using the Multi Gauge V3.0 program (Fujifilm, Tokyo, Japan).

Plasmids

pcDNA3_GRIM-19 was constructed to overexpress GRIM-19. The GRIM-19 gene was amplified by PCR with GRIM-19-specific primers (GRIM-19-*Hin*dIII-F, 5′-CCC AAGCTTACCATGGCGGCGTCAAAGGTG-3′ and GRIM-19-*Eco*R I-R, 5′-CGGAATTCTTACGTGTACCACATGAAGCCG-3′) using cDNAs that were reverse transcribed using random primers from RNA extracted from Huh7 cells. The PCR products were cut with *Hin*dIII and *Eco*RI and inserted into a pcDNA3 vector (Invitrogen) in frame. To evaluate the efficiency of transfection with foreign gene-encoding plasmid in Huh7 cells, pEGFP-C1_GRIM-19 was constructed. The GRIM-19 gene was amplified by PCR with GRIM-19-specific primers (GRIM-19-*Bgl*II, GGAAGATCTATGGCGGCGTCAAAGGTGAAG and GRIM-19-*Eco*RI-R) as described above. The PCR products were cut with *Bgl*II and *Eco*RI and inserted into the pEGFP-C1 vector (Clontech Laboratories, Mountain View, CA, USA) in frame. pcDNA3_EGFP was kindly provided by Dr. Sean B. Lee (Tulane University, New Orleans, LA, USA). pcDNA3.1-2xflag-SREBP1c was purchased from Addgene (Cambridge, MA, USA).

Transient Transfection

To investigate the effects of GRIM-19 on HCV replication and lipogenesis, Huh7 cells or Huh7 cells in which HCV replication occurs were transfected with various plasmids as described above using FUGENE HD (Promega, Madison, WI, USA) according to the manufacturer's protocol.

Real-time Quantitative Reverse Transcription-PCR (rqRT-PCR)

The levels of HCV RNA in Huh7 cells infected with HCVcc were evaluated to verify the anti-HCV effects of GRIM-19. Total RNA was extracted with TRIzol reagent (Invitrogen) and purified according to the manufacturer's recommendations. cDNA was synthesized from 2 μg of total RNA with primers specific for the HCV 5′UTR (HCV-5′UTR-R, 5′-ACCACAAGGCCTTTCGCAACCCAACGCTAC-3′) using ImProm-II reverse transcriptase (Promega). cDNA was then subjected to real-time, quantitative RT-PCR (rqRT-PCR) using primer pairs and a TaqMan probe targeting a region within the HCV 5′UTR as previously described (Kim et al., 2009). rqRT-PCR was performed using a LightCycler 480 Probes Master kit (Roche Applied Science) and a LightCycler 480 system (Roche Applied Science) according to the manufacturer's instructions. Endogenous mRNA levels of GRIM-19 and genes involved in lipid metabolism were also assessed using the LightCycler 480 Probes Master kit and the LightCycler 480 system with gene-specific primers and fluorescent probes recommended by Roche Universal Probe Library Design Center. The thermal conditions were designed using the Roche Universal Probe Library's thermocycling conditions following the manufacturer's instructions. Human β-actin was used as a reference gene. All fluorescence data were analyzed using LightCycler 4.0 software (Roche Applied Science), and C_t results were exported to Excel spreadsheets. The comparative C_t method was used for relative quantification and normalization.

Dual-luciferase Assay

Changes in HCV-internal ribosome entry site (IRES) activity were confirmed by a dual-luciferase assay. A dual-luciferase reporter construct was kindly provided by Dr. Jong-Won Oh (Yonsei University, Seoul, South Korea). It contains a CMV promoter-controlled *Renilla* luciferase reporter gene followed

by the HCV IRES-controlled firefly luciferase reporter gene. Huh7 cells infected with HCVcc were cotransfected with the dual-luciferase reporter construct and pcDNA3_GRIM-19 using fuGENE HD. At 48 h post-transfection, dual-luciferase assays were performed with a Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions.

Subcellular Fractionation

Huh7 cells infected with HCVcc were transfected with pcDNA3 or pcDNA3_GRIM-19. After 48 h, the cells were subjected to subcellular fractionation into nuclear and cytoplasmic fractions using an NE-PER kit (Pierce, Rockford, IL, USA) according to the manufacturer's recommendations.

Reverse Transcription-polymerase Chain Reaction (RT-PCR)

The mRNA levels of *bcl2* and *mmp2* were evaluated using RT-PCR. Total RNA extraction and cDNA synthesis using random primers were performed as described above. Gene amplification was performed with GoTaq Polymerase (Promega) and specific primer pairs for *bcl2* (bcl2-F, 5′-TCCCTCGCTGCACAAATACTC-3′, and bcl2-R, 5′-TTCTG CCCCTGCCAAATCT-3′) and *mmp2* (mmp2-F, 5′-CCACTGCC TTCGATACAC-3′, and mmp2-R, 5′-GAGCCACTCTCT GGAATCTTAAA-3′). The PCR program ran as follows: 10 min at 94°C; 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 45 s; followed by a final 10 min incubation at 72°C. The amplified products were separated on 1.5% agarose gels containing 0.5 mg/mL ethidium bromide. The nucleic acids were visualized under UV light using a Gel-Doc CQ system (Bio-Rad, Vienna, Austria), and the band densities of each gene were analyzed using the Multi Gauge V3.0 program with β-actin serving as a loading control.

Apoptosis Assays

Apoptosis was detected with Annexin V/propidium iodide (PI) staining (BD BioSciences) according to the manufacturer's instructions. In total, 10, 000 cells were counted by flow cytometry using a fluorescence-activated cell sorter (FACS, Becton-Dickinson, San Jose, CA, USA). The resulting data were analyzed using Summit 5.2 software (Beckman Coulter Inc., Miami, FL, USA).

Intracellular Lipid Droplet Quantification

Huh7 cells and HCVcc-infected Huh7 cells were treated with 100 μM oleic acid (OA) in serum-free DMEM containing 1% BSA at 24 h post-transfection with pcDNA3_GRIM-19 or pEGFP-C1_GRIM-19. Twenty-four hours later, the cells were subjected to Nile Red staining to evaluate the changes in intracellular lipid content. The cells were washed with ice-cold phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde for 5 min at room temperature. After being washed with PBS again, the cells were stained with Nile Red (0.5 μg/mL) and 4′,6-diamidino-2-phenyl-indole (DAPI, 1 μg/mL) (Sigma–Aldrich). After staining, intracellular LDs were quantified by measuring density of fluorescence with a microplate reader (Molecular Devices, Sunnyvale, CA, USA), and the results were normalized to the cellular DAPI content (Hur et al., 2012). The distribution of lipid in cells was observed under an LSM 510 inverted laser-scanning confocal microscope (Carl Zeiss, Jena, Germany).

Immunofluorescence Staining

Huh7 cells and HCVcc-infected Huh7 cells were fixed with 4% paraformaldehyde for 30 min, and permeabilized with PBS containing 0.2% Triton X-100 for 30 min at room temperature. After washing three times with PBS, the cells were treated with a blocking solution (PBS containing 1% BSA, 0.1% gelatin, and 5% goat serum) for 30 min at room temperature, incubated with primary antibody overnight at 4°C, and washed five times with PBS containing 1% BSA and 0.1% gelatin. The cells were further incubated with secondary antibodies (Molecular Probes, Eugene, OR, USA) for 2 h and washed five times with PBS. Nuclei were visualized using (DAPI) in PBS for 10 min. Stained slides were observed under an LSM 510 inverted laser-scanning confocal microscope (Carl Zeiss).

Statistical Analysis

All data are representative of a minimum of three independent experiments. The data are expressed as the mean ± SD or ±SEM. For comparison of multiple groups, one-way analysis of variance (ANOVA) with Tukey's *post hoc* test was used to define statistically significant differences among groups. For statistical comparisons between two groups, Student's *t*-test was used. The statistical significance of differences between groups is expressed by an asterisk ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

RESULTS

GRIM-19 Protein Levels Are Downregulated in HCV-infected Cells

First, we examined the expression of GRIM-19 protein in Huh7 cells infected with genotype 2a HCVcc. As shown in **Figure 1A**, GRIM-19 expression decreased in Huh7 cells infected with HCVcc. On the 3rd day of HCV infection, protein level of GRIM-19 was downregulated by approximately 20%. Moreover, GRIM-19 expression was decreased by approximately 50% on the 12th day of HCV infection. To confirm the downregulation of GRIM-19 in HCV-replicating cells, we assessed GRIM-19 expression in genotype 2a HCV SGR cells and FGR cells. In these cells, the GRIM-19 protein level was lower than that in Huh7 control cells (**Figure 1B**). Furthermore, in the genotype 3 HCV FGR cell line derived from Huh7.5.1 cells, the expression level of GRIM-19 was lower compared to that in parental Huh7.5.1 cells (**Figure 1C**). Interestingly, GRIM-19 mRNA levels were not altered in the cells with active HCV replication (**Figure 1D**). Next, we determined the protein levels of GRIM-19 in liver tissues from patients with chronic HCV infection. Compared to the liver tissues without viral hepatitis, HCV-infected livers expressed markedly lower levels of GRIM-19 protein (**Figure 1E**). These results suggest that HCV infection causes downregulation of GRIM-19 at the post-transcriptional level.

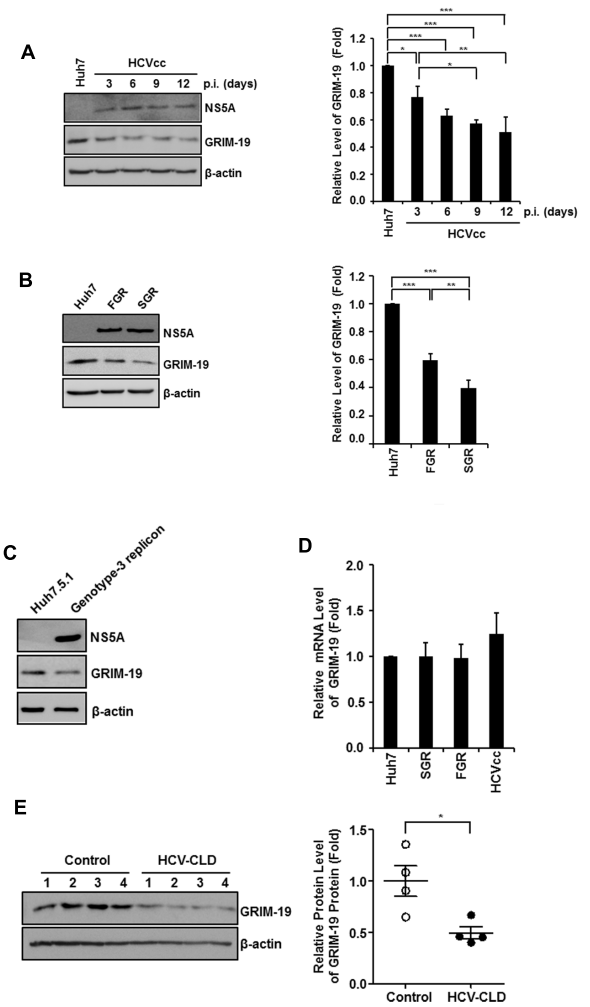


FIGURE 1 | Hepatitis C virus (HCV) infection and viral replication reduced GRIM-19 expression. (A,B) Protein levels of GRIM-19 were evaluated using western blot analysis in Huh7 cells infected with HCVcc at days 3, 6, 9, and 12 post-infection **(A)**, FGR cells **(B)**, and SGR cells **(B)**. The relative protein expression was normalized to β-actin as a reference. **(C)** The endogenous GRIM-19 level was assessed by western blot analysis in Huh7.5.1 cells or Huh7.5.1-derived HCV genotype 3 full genomic replicon cells. The relative protein expression was normalized to β-actin as a reference. **(D)** Relative levels of endogenous GRIM-19 mRNA in SGR cells, FGR cells, and HCVcc-infected Huh7 cells compared to that in Huh7 cells. β-actin was used as a reference gene. **(E)** Protein levels of GRIM-19 in liver tissues from patients with chronic liver diseases (CLD) caused by persistent HCV infection ($n = 4$) were analyzed by western blot analysis. Tissue lysates from liver without viral hepatitis were used as a control (normal; $n = 4$). β-actin was used as a loading control. The values of the GRIM-19 protein levels were expressed relative to the level in control tissues (right). All data represent the mean ± SEM ($n = 3$). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to control.

Ectopically Expressed GRIM-19 Reduces HCV RNA Replication

To investigate the roles of GRIM-19 in the HCV viral life cycle, we overexpressed GRIM-19 in HCVcc-infected or HCV replicon cells (**Figure 2**). GRIM-19 was ectopically expressed in the cells via transient transfection as described in section "Materials and Methods". Before evaluation of the effect of GRIM-19 overexpression on HCV replication, the transfection efficiency of Huh7 cells was checked using flow cytometry after transfection with enhanced green fluorescent protein (EGFP)-fused GRIM-19-encoding plasmids. As shown in **Figure 2B**, GFP fluorescence was detected in a high percentage of Huh7 cells transfected with EGFP-fused GRIM-19-encoding plasmids, even though the density of fluorescence among the cells was different. In the same transient transfection conditions, GRIM-19 was overexpressed in HCVcc-infected Huh7 cells. When Huh7 cells were infected with HCVcc at an MOI of 0.3, the level of intracellular HCV RNA was gradually increased until the 12th day after infection (**Figure 2G**). The level of intracellular HCV RNA on the 15th day was comparable to that on the 12th day (data not shown). These results indicate that the ratio of Huh7 cells infected with HCVcc to uninfected cells reached the highest level on approximately the 12th day after HCVcc infection at an MOI of 0.3. Therefore, to examine the effect of GRIM-19 overexpression on HCV replication, Huh7 cells infected with HCVcc were seeded on the 9th day post-infection, and the next day, the cells were transfected with GRIM-19 encoding plasmid. After 48 h, the levels of intracellular HCV RNA and protein were evaluated. As shown in **Figure 2C**, GRIM-19 overexpression resulted in an approximately 50% decrease in the levels of intracellular HCV RNA in Huh7 cells infected with HCVcc. Moreover, transient transfection with GRIM-19-encoding plasmid also reduced the protein level of HCV NS5A (**Figure 2C**, right). Ectopically expressed EGFP did not have an effect on the levels of either intracellular HCV RNA or HCV NS5A protein (**Figure 2C**). Interestingly, GRIM-19 transfected, HCVcc-infected cells secreted a much lower number of viral particles (**Figure 2D**, left). When Huh7 cells were re-infected with culture supernatants obtained from HCVcc-infected, GRIM-19-transfected cells, the levels of intracellular HCV RNA were lower (**Figure 2D**, right). Anti-HCV activity of GRIM-19 was also confirmed in FGR cells and SGR cells. As expected, GRIM-19 overexpression reduced the HCV RNA level to less than 50% in FGR cells and SGR cells (**Figures 2E** right, **F**). Additionally, ectopically expressed GRIM-19 reduced the protein level of HCV core in FGR cells (**Figure 2E**, right). Moreover, in the cells, EGFP overexpression did not have an effect on the levels of either intracellular HCV RNA or HCV core protein (**Figure 2E**). Furthermore, repeated transfection of Huh7 cells infected with HCVcc with GRIM-19 encoding plasmid resulted in additive inhibitory effects on HCV RNA replication (**Figure 2G**). In the first round of transfection with GRIM-19, the level of HCV RNA was 59% that of pcDNA3-transfected cells. After the fourth round of GRIM-19 transfection, the level of HCV RNA was 19% that of vehicle-transfected, HCVcc-infected Huh-7 cells. To confirm the suppressive function of GRIM-19 on HCV replication, we examined whether the inhibitory effect of GRIM-19 on HCV replication could be abolished

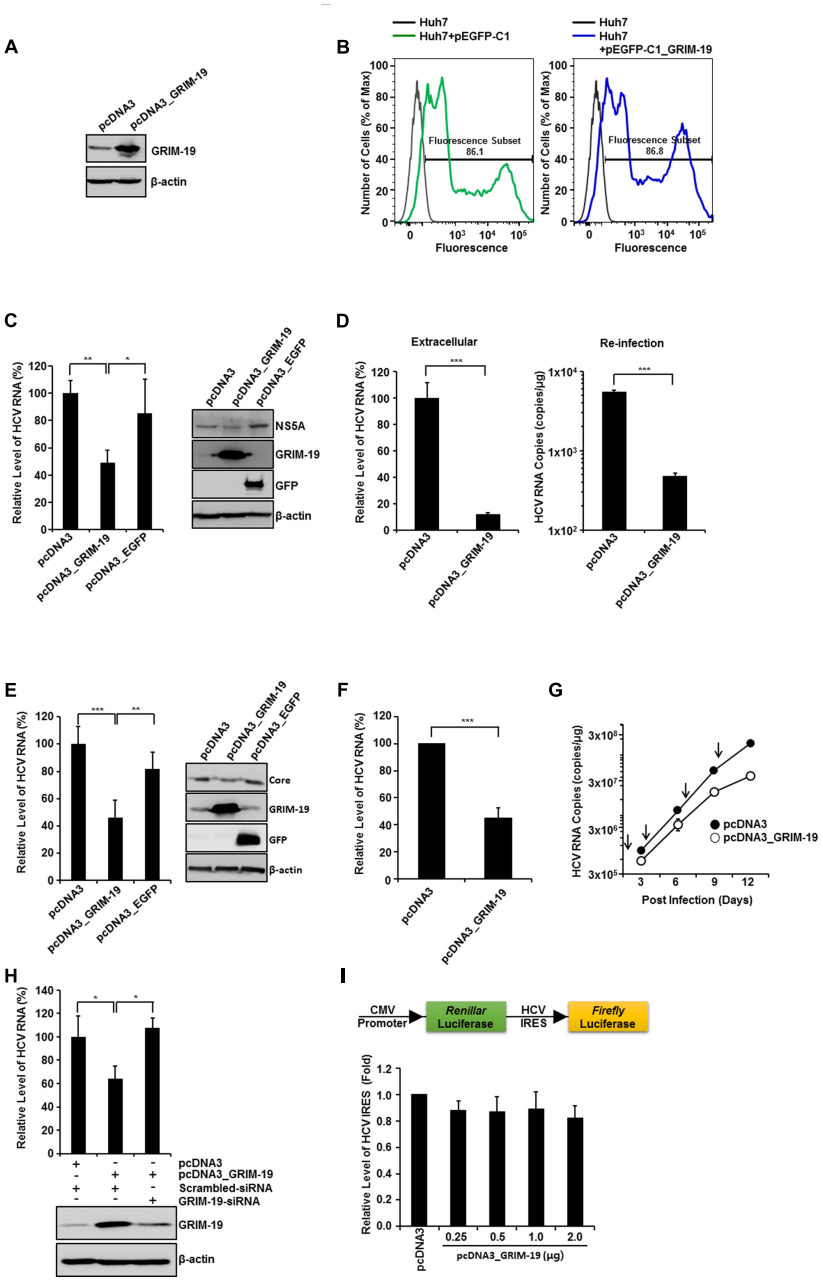


FIGURE 2 | Anti-HCV activity of ectopically expressed GRIM-19. (A) Overexpression of GRIM-19 was induced by transfection with pcDNA3_GRIM-19 in Huh7 cells. At 48 h post-transfection, GRIM-19 overexpression was detected by western blot analysis. **(B)** Huh7 cells were transfected with pEGFP-C1 or pEGFP-C1-GRIM-19. After 48 h, the cells were subjected to flow cytometry to evaluate the transfection efficiency. **(C)** HCVcc-infected Huh7 cells at day 9 post-infection were transfected with pcDNA3, pcDNA3_GRIM-19 or pcDNA3_EGFP. After 48 h, the intracellular levels of HCV RNA were evaluated by rqRT PCR (left), and the protein level of HCV NS5A was detected by western blot analysis (right). The values of the HCV RNA levels were expressed relative to the level in cells transfected with pcDNA3. **(D)** At day 9 post-infection, HCVcc-infected Huh7 cells were transfected with pcDNA3 or pcDNA3_GRIM-19. After 48 h, the culture media was used to evaluate the extracellular level of HCV RNA (left) as well as for re-infection. After another 48 h, the HCV RNA levels in re-infected Huh7 cells were evaluated by rqRT PCR. **(E)** Levels of HCV RNA and HCV core protein in FGR cells were analyzed after transfection with pcDNA3, pcDNA3_GRIM-19 or pcDNA3_EGFP as in **(C)**. **(F)** The values of the HCV RNA levels are analyzed in SGR cells transfected with pcDNA3 or pcDNA3_GRIM-19 as in **(C)**. **(G)** To examine the inhibitory effect of repeated transfection with pcDNA3_GRIM-19, Huh7 cells were infected with HCVcc at an MOI of 0.3. After 24 h, cells were transfected with pcDNA3 or pcDNA3_GRIM-19. The cells were maintained at 3-day intervals and transfected with pcDNA3 or pcDNA3_GRIM-19 repeatedly. At 3, 6, 9, and 12 days post-infection, cells were harvested, and HCV RNA levels were evaluated. **(H)** At day 9 post-infection, HCVcc-infected Huh7 cells were transfected with pcDNA3 or pcDNA3_GRIM-19. After 24 h, the cells were transfected with scrambled siRNA or GRIM-19-siRNA. Changes in GRIM-19 levels were assessed by western blot analysis 48 h post-transfection with siRNAs (bottom). In addition, HCV RNA levels were evaluated by rqRT-PCR (top). **(I)** HCV-IRES activity in Huh7 transfected with pcDNA3 or pcDNA3_GRIM-19. The data represent the means ± SEM (n = 3). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to control.

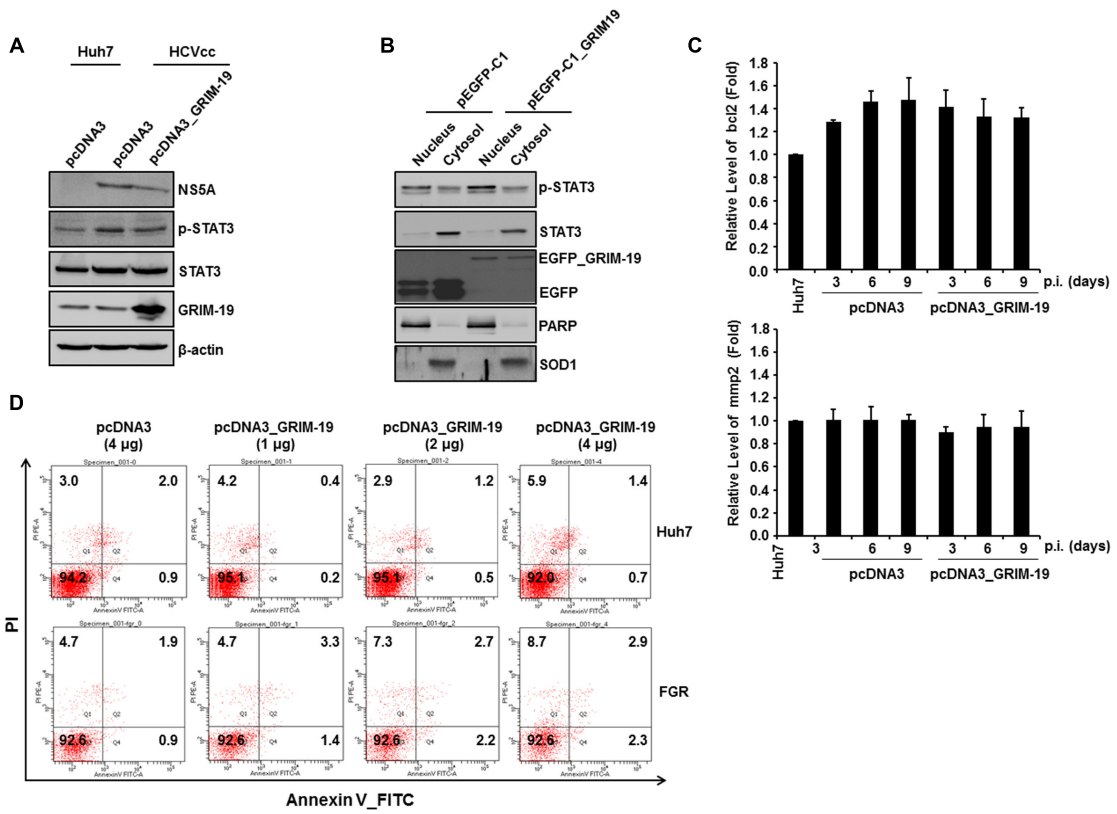


FIGURE 3 | GRIM-19 overexpression did not alter STAT3 transcriptional activity. (A) The expression and phosphorylation of STAT3 after transfection with pcDNA3 or pcDNA3_GRIM-19 were evaluated by western blot analysis. β-actin was used as an internal control for loading. **(B)** To evaluate subcellular localization of phospho-STAT3, STAT3, and GRIM-19, HCVcc-infected Huh7 cells were transfected with pEGFP-C1 or pEGFP-C1_GRIM-19. After 48 h, the cells were subjected to subcellular fractionation followed by western blot analysis. PARP and SOD1 were used as markers for the nucleus and cytosol. **(C)** The mRNA levels of *bcl2* and *mmp2* were assessed by RT-PCR with a gene-specific primer set. The mRNA expression was normalized to β-actin as a reference and expressed relative to the density of Huh7 cells. Values represent means ± SD. **(D)** Examination of the induction of apoptosis by overexpression of GRIM-19 in Huh7 cells or FGR cells. The cells were transfected with pcDNA3 or pcDNA3_GRIM-19. After 48 h, apoptosis was determined by a Annexin V/PI staining.

by siRNA against GRIM-19. In HCVcc-infected, GRIM-19-overexpressing Huh7 cells, transfection with GRIM-19 siRNA abrogated the suppressive effect of GRIM-19 on HCV replication (**Figure 2H**). Collectively, these results suggest that GRIM-19 overexpression inhibits HCV.

Viral Internal Ribosome Entry Site (IRES)-Mediated Translation of the HCV Genome Is Not Altered by GRIM-19

Next, we investigated whether GRIM-19 overexpression had an effect on HCV-IRES activity. Huh7 cells were transfected with GRIM-19-encoding plasmid and a dual-luciferase reporter construct, allowing cap-dependent expression of *Renilla* luciferase and HCV IRES-dependent translation of firefly luciferase. As shown in **Figure 2I**, forced expression of GRIM-19 did not alter HCV IRES activity. These results suggest that the suppressive effect of GRIM-19 on HCV RNA replication is not caused by alteration of HCV-IRES activity.

GRIM-19 Overexpression Does Not Alter the Subcellular Localization or the Transcriptional Activity of Phosphorylated STAT3 in HCVcc-infected Cells

As described above, one of the host factors that interact with HCV proteins is STAT3. Because GRIM-19 is known to interact with phosphorylated STAT3 and transport it out of the nucleus (Shulga and Pastorino, 2012), we investigated whether GRIM-19 overexpression has an effect on STAT3 activation in HCVcc-infected Huh7 cells. First, we aimed to confirm the amount and subcellular localization of phosphorylated STAT3 in Huh7 cells infected with HCVcc after transfection with a GRIM-19-encoding plasmid. The level of phosphorylated STAT3 increased after infection with HCVcc in Huh7 cells (**Figure 3A**); however, overexpression of GRIM-19 did not reduce the increased phosphorylation of STAT3 (**Figure 3A**). Furthermore, GRIM-19 overexpression did not induce translocation of phosphorylated STAT3 (**Figure 3B**). Follow-up experiments showed that the expression levels of *bcl2* and *mmp2*, which are induced

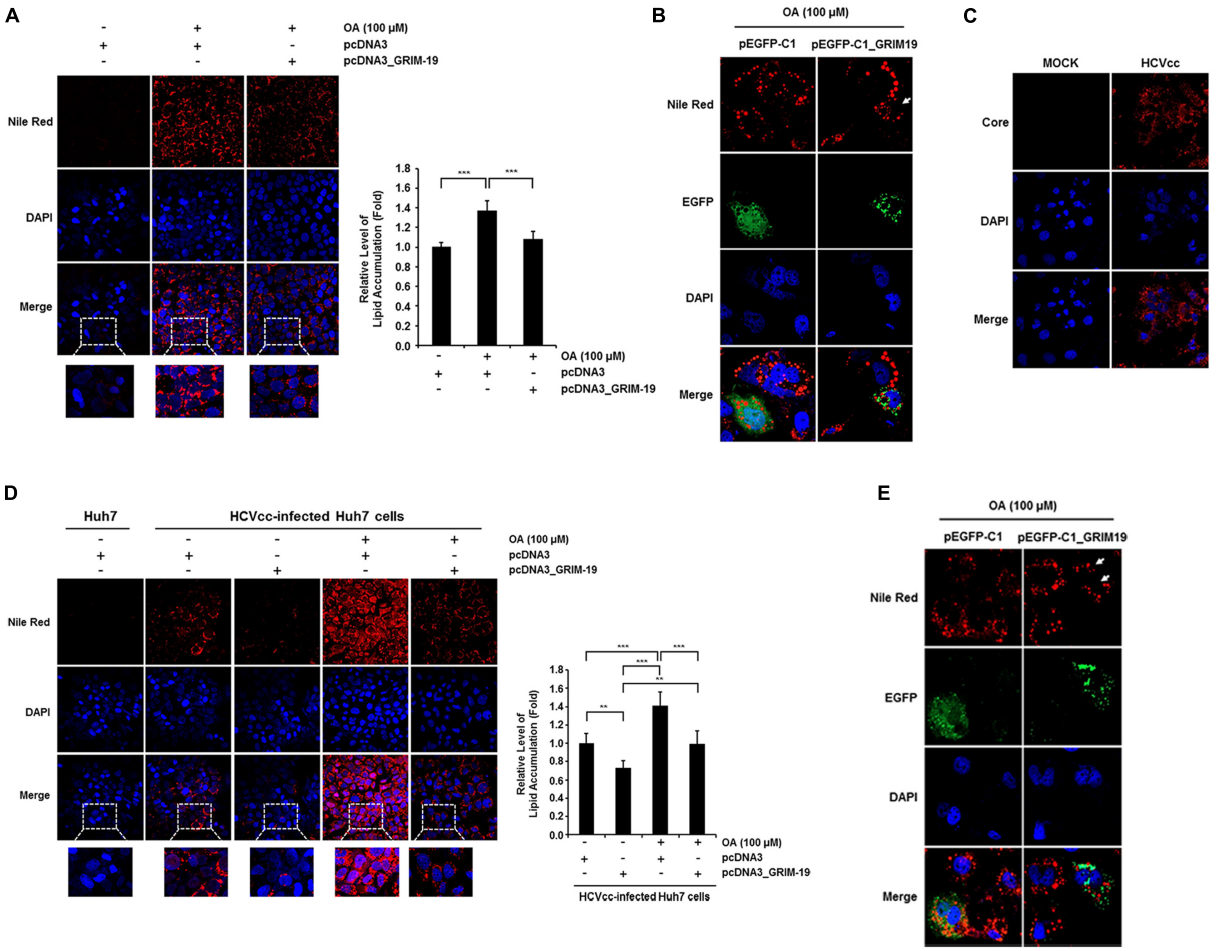


FIGURE 4 | GRIM-19 overexpression inhibited lipid accumulation. (A) Huh7 cells were transfected with pcDNA3 or pcDNA3_GRIM-19. After 24 h, cells were treated with 100 μM OA. After an additional 24 h, intracellular lipid accumulation was visualized by Nile Red staining and observed under a confocal microscope (left). Red, lipid droplets (LDs); blue, DAPI; original magnification 200×. Fluorescence densities of Nile Red were quantified using a microplate reader and normalized to the cellular DAPI content. The values of intracellular lipid level were expressed relative to the level in the cells transfected with pcDNA3 and without OA treatment (right). The data represent the mean ± SEM ($n = 3$). $^{**}P < 0.01$, $^{***}P < 0.001$ compared to control. **(B)** Huh7 cells were transfected with pEGFP-C1 or pEGFP-C1-GRIM-19. After 24 h, cells were treated with 100 μM OA. After an additional 24 h, intracellular lipid accumulation was assessed by Nile Red staining. Red, LDs; Green, EGFP or EGFP-fused GRIM-19; blue, DAPI; original magnification 800×. **(C)** HCV core was immunostained in Huh7 cells infected with HCVcc at day 9 post-infection. Red, HCV core; blue, DAPI; original magnification 400×. **(D)** Changes of intracellular lipid accumulation by GRIM-19 overexpression was assessed in Huh7 cells infected with HCVcc at day 9 post-infection as in **(A)**. **(E)** Smaller size and number of LDs in GRIM-19 overexpressing cells were visualized in HCVcc infected Huh7 cells as in **(B).**

by transcriptional activation of STAT3, were not altered by overexpression of GRIM-19 (**Figure 3C**). Furthermore, GRIM19 overexpression did not induce sufficient apoptosis to show anti-HCV activity in HCV-replicating Huh7 cells (**Figure 3D**). These results suggest that the inhibitory effect of GRIM-19 on HCV replication is not associated with altered STAT3 activation in HCV-infected cells.

GRIM-19 Attenuates Intracellular Lipid Accumulation

The level of HCV RNA was significantly downregulated by transfection with GRIM-19-encoding plasmids in SGR cells (**Figure 2F**). This suggests that the anti-viral function of GRIM-19 may be closely associated with HCV replication or viral RC formation. As shown in **Figure 2I**, GRIM-19 overexpression does not restrict viral translation. In addition, the dependence of HCV replication and viral RC formation on intracellular lipid accumulation is well known (Kapadia and Chisari, 2005; Mankouri et al., 2010; Pisonero-Vaquero et al., 2014; Akil et al., 2016). Therefore, we further investigated whether GRIM-19 overexpression affects intracellular lipid accumulation. Huh7 cells and HCVcc-infected Huh7 cells were transfected with GRIM-19-encoding plasmids. In both types of cells, treatment with OA increased intracellular lipid levels by up to approximately 140% (**Figures 4A,D**). However, in cells overexpressing GRIM-19, the levels of intracellular lipid accumulation after treatment with OA were comparable to that

of untreated cells (**Figures 4A,D**). Moreover, the number and size of LDs after OA treatment were smaller in the cells transfected with EGFP-fused GRIM-19-encoding plasmids compared to that in untransfected cells (**Figures 4B,E**). Furthermore, as shown in **Figure 4C**, HCV core protein was detected in most Huh7 cells infected with HCVcc on the 9th day post-infection, and in the cells, the amount of intracellular lipids increased without OA treatment, as previously reported (**Figure 4D**) (McRae et al., 2015; Akil et al., 2016). Additionally, overexpression of GRIM-19 significantly reduced the level of intracellular lipid accumulation caused by HCV infection (**Figure 4D**). These results may indicate that GRIM-19 can ameliorate intracellular lipid accumulation in hepatocytes.

GRIM-19 Overexpression Downregulates the Expression Levels of SREBP1-c and Its Target Genes

Next, we investigated the mRNA level of transcription factors involved in regulating the intracellular lipid level after transfection with GRIM-19-encoding plasmids. We examined the expression of the following transcription factors: (i) SREBP-1c, known to induce *de novo* lipogenesis to generate free fatty acid (FFA) (Sanders and Griffin, 2016); (ii) peroxisome proliferator-activated receptor α (PPARα), required for mitochondrial, peroxisomal, and microsomal FFA oxidation (Memon et al., 2000); and (iii) peroxisome proliferator-activated receptor γ (PPARγ), which is known to contribute to FFA uptake (Ahmadian et al., 2013). Among these three transcription factors, only SREBP-1c expression levels were significantly upregulated in Huh7 cells treated with 100 μM OA and downregulated by transfection with GRIM-19-encoding plasmids before OA treatment (**Figures 5A** left, **C**). Interestingly, the mRNA levels of SREBP-1c target genes involved in triglyceride biosynthesis such as FAS, SCD, and ACC were upregulated following treatment with OA and downregulated as a result of GRIM-19 overexpression induced prior to OA treatment (**Figure 5B**). Moreover, increased ACC and FAS protein levels resulting from OA treatment were downregulated by ectopically expressed GRIM-19 (**Figure 5C**). In both FGR cells and Huh7 cells infected with HCVcc, the expression levels of SREBP-1c, FAS, and ACC were remarkably downregulated by GRIM-19 overexpression (**Figures 5D,E**). Unexpectedly, the protein levels of SCD were not significantly altered by GRIM-19 overexpression in the cells (**Figures 5C–E**). Next, we further determined the effect of GRIM-19 on the expression levels of other enzymes involved in lipid metabolism. Interestingly, the expression levels of diacylglycerol acyltransferase-1 (DGAT-1) and diacylglycerol acyltransferase-2 (DGAT-2), which catalyze the final step in triglyceride biosynthesis, were not significantly affected by GRIM-19 overexpression (**Figure 5F**). Likewise, microsomal triglyceride transfer protein (MTP), which is involved in the assembly/secretion of very low density lipoproteins (VLDL), was not upregulated by OA treatment even though overexpression of GRIM-19 decreased the MTP mRNA level (**Figure 5F**). Taken together, these results demonstrate that GRIM-19 may ameliorate intracellular lipid accumulation by regulating the expression of SREBP-1c and its target genes.

GRIM-19 Restricts HCV Replication through Downregulation of SREBP-1c

To confirm the suppressive effect of GRIM-19 on intracellular lipid accumulation, we examined whether the restrictive effect of GRIM-19 on HCV replication could be abolished by OA treatment or normalization of SREBP-1c expression. As shown in **Figure 6A**, OA treatment in the absence of GRIM-19 transfection increased the HCV RNA titer in HCVcc-infected Huh7 cells. Moreover, in HCVcc-infected, GRIM-19-overexpressing Huh7 cells, downregulation of HCV RNA was restored by OA treatment (**Figure 6A**). In the same manner, SREBP-1c overexpression increased the level of HCV RNA and abolished the restrictive effect of GRIM-19 on HCV replication (**Figure 6B**). These results suggest that the inhibitory effect of GRIM-19 overexpression on HCV replication is mediated by downregulation of SREBP-1c.

DISCUSSION

In this study, we uncovered a new biological function of GRIM-19 in lipid metabolism, as summarized in **Figure 7**. HCV replication caused the GRIM19 protein level to decrease. However, restoration of the downregulated level of GRIM-19 by transient transfection with a GRIM-19-encoding plasmid restricted HCV replication. Interestingly, GRIM-19 overexpression downregulated the expression of SREBP-1c and its target genes, resulting in abrogation of the intracellular lipid accumulation induced by HCV replication. These results suggest that GRIM-19 can be thought of as a host factor that restricts HCV replication.

It is known that HCV exploits host lipid architectures and molecules involved in lipid metabolism for its efficient replication and propagation (Miyanari et al., 2007; Negro and Sanyal, 2009; Syed et al., 2010). For example, HCV genome replication, in common with other positive-strand RNA viruses, occurs within a "membranous web" derived from intracellular vesicles (Egger et al., 2002; Miyanari et al., 2007). Additionally, it has been reported that LDs act as a platform for HCV replication and assembly (Miyanari et al., 2007). HCV core protein recruits HCV RNA, non-structural proteins, and replication complexes to LD-associated membranes. Thus, this recruitment is critical for infectious virus particle production (Miyanari et al., 2007). For these reasons, HCV induces intracellular lipid accumulation to optimize the cellular environment for persistent infection (Kapadia and Chisari, 2005; Mankouri et al., 2010; Akil et al., 2016). As one of the strategies to increase intracellular lipids, HCV activates SREBPs (Waris et al., 2007; Negro and Sanyal, 2009; Syed et al., 2010). Recent research has demonstrated that HCV NS4B, NS5A, and core protein may activate SREBP-1c and its target genes, resulting in enhanced fatty acid biosynthesis (Park et al., 2009; Syed et al., 2010; Xiang et al., 2010; Garcia-Mediavilla et al., 2012). Moreover, inhibiting or silencing SREBP by treatment with 25-hydroxycholesterol inhibits HCV

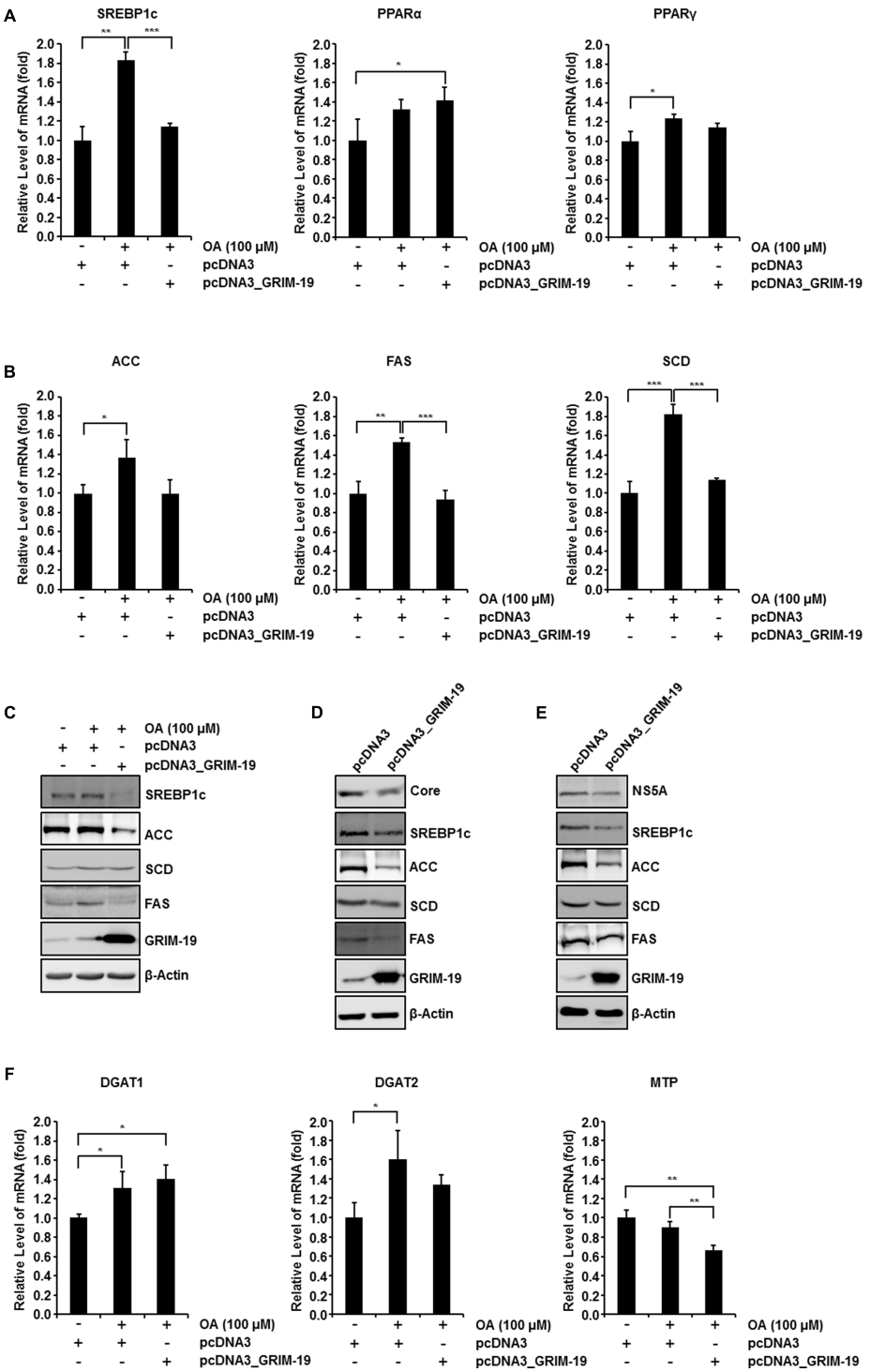


FIGURE 5 | Effect of GRIM-19 overexpression on the expression levels of genes involved in lipid metabolism. (A) Examination of the mRNA levels of three transcription factors that regulate the intracellular lipid level in Huh7 cells treated with OA or transfected with pcDNA3 or pcDNA3_GRIM-19. The mRNA expression was normalized to β-actin as a reference and the values of mRNA level were expressed relative to the level in cells transfected with pcDNA3 and without OA treatment. **(B)** mRNA levels of target genes of SREBP-1c were analyzed as in **(A). (C)** Protein levels of SREBP-1c and its target genes in Huh7 cells treated as in **(A)** were analyzed using western blot analysis. β-actin was used as an internal control for loading. **(D,E)** Protein levels of SREBP-1c and its target genes after transfection with pcDNA3 or pcDNA3_GRIM-19 in FGR cells **(D)** and Huh7 cells infected with HCVcc at day 9 post-infection **(E)**. β-actin was used as an internal control for loading. **(F)** mRNA levels of DGAT1, DGAT-2, and MTP were analyzed as in **(A)**. The data represent the mean ± SEM ($n = 3$). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to control.

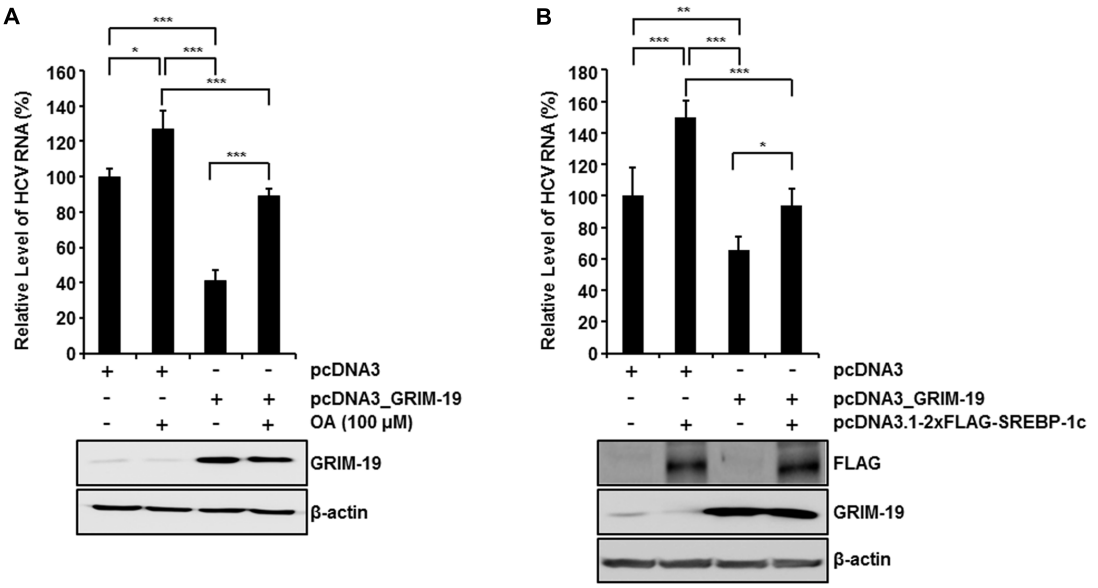


FIGURE 6 | Restoration of GRIM-19 inhibition of HCV replication. At day 9 post-infection, HCVcc-infected Huh7 cells were transfected with pcDNA3 or pcDNA3_GRIM-19. After 24 h, the cells were treated with OA **(A)** or transfected with SREBP-1c-overexpressing plasmids **(B)**. After an additional 48 h, intracellular levels of HCV RNA were evaluated by rqRT PCR (top), and the overexpressed protein levels was detected by western blot analysis (bottom). The values of the HCV RNA levels were expressed relative to the level in cells transfected with pcDNA3. The data represent the mean ± SEM (n = 3). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to control.

replication (Su et al., 2002; Yang et al., 2008; Li et al., 2013). In the present study, we showed that GRIM-19 overexpression impeded the lipid accumulation induced by OA treatment and HCV infection through downregulation of SREBP-1c expression. Furthermore, the expression of SREBP-1c target genes, such as ACC and FAS but not SCD was markedly downregulated by GRIM-19 overexpression. Luyimbazi et al. (2010) demonstrated that translation of SCD is regulated by eukaryotic initiation factor 4E (eIF4E), suggesting that mTOR may regulate SCD through the mTOR/4E-BP1/eIF4E axis. More recently, it was reported that eIF4E could be activated in HCV-replicating cells for efficient viral translation (George et al., 2012; Licursi et al., 2012; Panda et al., 2014). Based on these reports, we speculate that eIF4E activation by HCV could be one of the reasons that mRNA levels of SCD are decreased by GRIM-19 overexpression but the protein levels of SCD are not affected under the same conditions. However, future investigations are needed to understand the exact mechanisms of a sustained level of SCD protein after transfection with GRIM-19-encoding plasmids. Despite the fact that SCD expression was not modulated by GRIM-19 overexpression, these findings suggest that GRIM-19 may abrogate intracellular lipid accumulation by inhibiting SREBP-1c activation and that these inhibitory effects could restrict HCV replication.

Signal transducer and activator of transcription 3, a major oncogenic transcription factor involved in cancer development and progression, is regulated by GRIM19 in some tumor cell lines (Alchanati et al., 2006; Kalakonda et al., 2007), but little is known regarding their relationships in viral infections. For this reason, we explored a specific functional link between GRIM19 and STAT3 in HCV infection. However, our results showed that the inhibitory effect of GRIM-19 on HCV replication was not associated with altered STAT3 activation in HCV-infected cells. Moreover, the precise mechanism by which GRIM-19 downregulates SREBP-1c gene expression remains

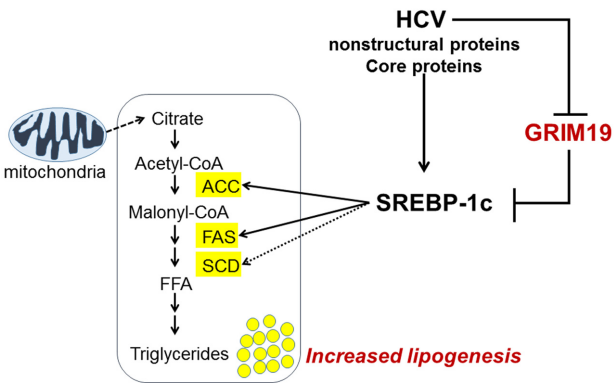


FIGURE 7 | A proposed model of the mechanism underlying the anti-HCV activity of GRIM-19. After HCV infection, lipid biosynthesis is upregulated via SREBP-1c activation. SREBP-1c is a transcription factor that enhances the expression of genes involved in fatty acid biosynthesis such as ACC, FAS, and SCD, causing increased levels of triglycerides that help create an appropriate microenvironment for persistent HCV infection. Conversely, GRIM-19 function is downregulated through the reduction of its protein levels in HCV-replicating cells. However, restoration of GRIM-19 expression can disrupt this suitable microenvironment by downregulating SREBP-1c expression.

unclear. Hence, further investigation is warranted concerning the functional links among GRIM-19, STAT3, and SREBP-1c in HCV infection. Interestingly, a recent study showed that leptin-induced STAT3 downregulates SREBP-1c expression in hepatic stellate cells (Zhang et al., 2013). Therefore, it appears that STAT3 activity is not strongly associated with GRIM-19 inhibition of SREBP-1c expression.

To date, the role of p53 in lipid metabolism is still uncertain; however, p53 and its target genes may be anti-viral host factors in HCV pathogenesis because they are also impaired by HCV infection (Nishimura et al., 2009). Recently, Zhou et al. (2011) demonstrated that the p53 tumor suppressor was regulated by GRIM-19 expression. They showed that GRIM-19 helps to stabilize p53 by interacting with E6 and E6AP proteins and inducing ubiquitination and degradation of E6AP, resulting in the promotion of apoptosis in a cervical cancer cell line (Zhou et al., 2011). In contrast, another study by Ruedo-Rincon et al. (2015) demonstrated that p53 could decrease SCD expression by repressing SREBP-1c. Based on these results, we speculate that p53 is involved in the negative regulation of SREBP-1c expression by GRIM-19. However, future investigations are needed to understand the exact mechanisms of the interaction between p53 and GRIM-19 during SREBP1-c expression and lipid metabolism.

CONCLUSION

Our data reveal a previously unknown role of GRIM-19 in lipid metabolism during HCV pathogenesis. GRIM-19 overexpression abrogated intracellular lipid accumulation through downregulation of SREBP-1c and its target genes, resulting in restriction of HCV replication. These results provide valuable information regarding GRIM-19, a host factor involved in HCV replication. Furthermore, this new knowledge regarding GRIM-19 may facilitate new strategies against diseases related to lipid metabolic disorders.

AUTHOR CONTRIBUTIONS

SY, E-CS, and J-HK designed the concepts of this study. J-HK, PS, EL, and DP carried out the experiments. SY, E-CS, J-HK, PS, WH, and MW discussed and interpreted the results. J-HK wrote the manuscripts. SY supervised the experiment and project.

FUNDING

This study was partially supported by Research Fund of Seoul St. Mary's Hospital, the Catholic University of Korea and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2015R1C1A1A02037212). This research was also co-supported by the Global High-tech Biomedicine Technology Development Program of the National Research Foundation (NRF) and Korea Health Industry Development Institute (KHIDI) funded by the Korean government (MSIP&MOHW) (No. HI15C3516).

ACKNOWLEDGMENTS

We thank Dr. Takaji Wakita (National Institute of Infectious Diseases, Tokyo, Japan) for use of pJFH1, pSGR-JFH1, and pFGR-JFH1, Dr. Francis Chisari (The Scripps Research Institute, CA) for providing Huh7.5.1 cells, Dr. Jane C. Moores (The Regent of the University of California, Oakland, CA, USA) for providing Huh7 cells and Huh7.5 cells, Dr. Sung K. Jang (Pohang University of Science and Technology, Pohang, Kyungbuk, South Korea) for supporting evaluation of GRIM-19 expression in HCVgenotype-3 replicon cells, Dr. Sean B. Lee (Tulane University, New Orleans, LA, USA) for use of pcDNA3_EGFP, and Dr. Jong-Won Oh (Yonsei University, Seoul, South Korea) for providing a dual-luciferase reporter construct.

REFERENCES

Ahmadian, M., Suh, J. M., Hah, N., Liddle, C., Atkins, A. R., Downes, M., et al. (2013). PPARgamma signaling and metabolism: the good, the bad and the future. *Nat. Med.* 19, 557–566. doi: 10.1038/nm.3159

Aizaki, H., Lee, K. J., Sung, V. M., Ishiko, H., and Lai, M. M. (2004). Characterization of the hepatitis C virus RNA replication complex associated with lipid rafts. *Virology* 324, 450–461. doi: 10.1016/j.virol.2004.03.034

Akil, A., Wedeh, G., Zahid Mustafa, M., and Gassama-Diagne, A. (2016). SUMO1 depletion prevents lipid droplet accumulation and HCV replication. *Arch. Virol.* 161, 141–148. doi: 10.1007/s00705-015-2628-3

Alchanati, I., Nallar, S. C., Sun, P., Gao, L., Hu, J., Stein, A., et al. (2006). A proteomic analysis reveals the loss of expression of the cell death regulatory gene GRIM-19 in human renal cell carcinomas. *Oncogene* 25, 7138–7147. doi: 10.1038/sj.onc.1209708

Angell, J. E., Lindner, D. J., Shapiro, P. S., Hofmann, E. R., and Kalvakolanu, D. V. (2000). Identification of GRIM-19, a novel cell death-regulatory gene induced by the interferon-beta and retinoic acid combination, using a genetic approach. *J. Biol. Chem.* 275, 33416–33426. doi: 10.1074/jbc.M003929200

Aslam, R., Raza, S. M., Naeemi, H., Mubarak, B., Afzal, N., and Khaliq, S. (2016). SOCS3 mRNA expression and polymorphisms as pretreatment predictor of response to HCV genotype 3a IFN-based treatment. *Springerplus* 5, 1826. doi: 10.1186/s40064-016-3506-5

Choi, J. E., Kwon, J. H., Kim, J. H., Hur, W., Sung, P. S., Choi, S. W., et al. (2015). Suppression of dual specificity phosphatase I expression inhibits hepatitis C virus replication. *PLoS ONE* 10:e0119172. doi: 10.1371/journal.pone.0119172

Date, T., Miyamoto, M., Kato, T., Morikawa, K., Murayama, A., Akazawa, D., et al. (2007). An infectious and selectable full-length replicon system with hepatitis C virus JFH-1 strain. *Hepatol. Res.* 37, 433–443. doi: 10.1111/j.1872-034X.2007.00056.x

Egger, D., Wolk, B., Gosert, R., Bianchi, L., Blum, H. E., Moradpour, D., et al. (2002). Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J. Virol.* 76, 5974–5984. doi: 10.1128/JVI.76.12.5974-5984.2002

El-Saadany, S., Ziada, D. H., El Bassat, H., Farrag, W., El-Serogy, H., Eid, M., et al. (2013). The role of hepatic expression of STAT1, SOCS3 and PIAS1 in the response of chronic hepatitis C patients to therapy. *Can. J. Gastroenterol.* 27, e13–e17. doi: 10.1155/2013/562765

Gale, M. Jr., and Foy, E. M. (2005). Evasion of intracellular host defence by hepatitis C virus. *Nature* 436, 939–945. doi: 10.1038/nature04078

Garcia-Mediavilla, M. V., Pisonero-Vaquero, S., Lima-Cabello, E., Benedicto, I., Majano, P. L., Jorquera, F., et al. (2012). Liver X receptor alpha-mediated regulation of lipogenesis by core and NS5A proteins contributes to HCV-induced liver steatosis and HCV replication. *Lab. Invest.* 92, 1191–1202. doi: 10.1038/labinvest.2012.88

George, A., Panda, S., Kudmulwar, D., Chhatbar, S. P., Nayak, S. C., and Krishnan, H. H. (2012). Hepatitis C virus NS5A binds to the mRNA cap-binding eukaryotic translation initiation 4F (eIF4F) complex and up-regulates host translation initiation machinery through eIF4E-binding protein 1 inactivation. *J. Biol. Chem.* 287, 5042–5058. doi: 10.1074/jbc.M111.308916

Hur, W., Kim, S. W., Lee, Y. K., Choi, J. E., Hong, S. W., Song, M. J., et al. (2012). Oleuropein reduces free fatty acid-induced lipogenesis via lowered extracellular signal-regulated kinase activation in hepatocytes. *Nutr. Res.* 32, 778–786. doi: 10.1016/j.nutres.2012.06.017

Kalakonda, S., Nallar, S. C., Lindner, D. J., Hu, J., Reddy, S. P., and Kalvakolanu, D. V. (2007). Tumor-suppressive activity of the cell death activator GRIM-19 on a constitutively active signal transducer and activator of transcription 3. *Cancer Res.* 67, 6212–6220. doi: 10.1158/0008-5472.CAN-07-0031

Kalvakolanu, D. V., Nallar, S. C., and Kalakonda, S. (2010). Cytokine-induced tumor suppressors: a GRIM story. *Cytokine* 52, 128–142. doi: 10.1016/j.cyto.2010.03.009

Kapadia, S. B., and Chisari, F. V. (2005). Hepatitis C virus RNA replication is regulated by host geranylgeranylation and fatty acids. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2561–2566. doi: 10.1073/pnas.0409834102

Kim, S. J., Kim, J. H., Sun, J. M., Kim, M. G., and Oh, J. W. (2009). Suppression of hepatitis C virus replication by protein kinase C-related kinase 2 inhibitors that block phosphorylation of viral RNA polymerase. *J. Viral. Hepat.* 16, 697–704. doi: 10.1111/j.1365-2893.2009.01108.x

Kong, L., Li, S., Yu, X., Fang, X., Xu, A., Huang, M., et al. (2016). Hepatitis C virus and its protein NS4B activate the cancer-related STAT3 pathway via the endoplasmic reticulum overload response. *Arch. Virol.* 161, 2149–2159. doi: 10.1007/s00705-016-2892-x

Li, M., Li, Z., Liang, C., Han, C., Huang, W., and Sun, F. (2014). Upregulation of GRIM-19 suppresses the growth of oral squamous cell carcinoma in vitro and in vivo. *Oncol. Rep.* 32, 2183–2190. doi: 10.3892/or.2014.3423

Li, Q., Brass, A. L., Ng, A., Hu, Z., Xavier, R. J., Liang, T. J., et al. (2009). A genome-wide genetic screen for host factors required for hepatitis C virus propagation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16410–16415. doi: 10.1073/pnas.0907439106

Li, Q., Pene, V., Krishnamurthy, S., Cha, H., and Liang, T. J. (2013). Hepatitis C virus infection activates an innate pathway involving IKK-alpha in lipogenesis and viral assembly. *Nat. Med.* 19, 722–729. doi: 10.1038/nm.3190

Li, Q., Zhang, Y. Y., Chiu, S., Hu, Z., Lan, K. H., Cha, H., et al. (2014). Integrative functional genomics of hepatitis C virus infection identifies host dependencies in complete viral replication cycle. *PLoS Pathog.* 10:e1004163. doi: 10.1371/journal.ppat.1004163

Licursi, M., Komatsu, Y., Pongnopparat, T., and Hirasawa, K. (2012). Promotion of viral internal ribosomal entry site-mediated translation under amino acid starvation. *J. Gen. Virol.* 93, 951–962. doi: 10.1099/vir.0.040386-0

Lindenbach, B. D., and Rice, C. M. (2005). Unravelling hepatitis C virus replication from genome to function. *Nature* 436, 933–938. doi: 10.1038/nature04077

Liu, S., Zhang, W., Liu, K., Wang, Y., Ji, B., and Liu, Y. (2014). Synergistic effects of co-expression plasmidbased ADAM10-specific siRNA and GRIM-19 on hepatocellular carcinoma in vitro and in vivo. *Oncol. Rep.* 32, 2501–2510. doi: 10.3892/or.2014.3503

Luyimbazi, D., Akcakanat, A., Mcauliffe, P. F., Zhang, L., Singh, G., Gonzalez-Angulo, A. M., et al. (2010). Rapamycin regulates stearoyl CoA desaturase 1 expression in breast cancer. *Mol. Cancer Ther.* 9, 2770–2784. doi: 10.1158/1535-7163.MCT-09-0980

Mankouri, J., Tedbury, P. R., Gretton, S., Hughes, M. E., Griffin, S. D., Dallas, M. L., et al. (2010). Enhanced hepatitis C virus genome replication and lipid accumulation mediated by inhibition of AMP-activated protein kinase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11549–11554. doi: 10.1073/pnas.0912426107

McCartney, E. M., Helbig, K. J., Narayana, S. K., Eyre, N. S., Aloia, A. L., and Beard, M. R. (2013). Signal transducer and activator of transcription 3 is a proviral host factor for hepatitis C virus. *Hepatology* 58, 1558–1568. doi: 10.1002/hep.26496

McRae, S., Iqbal, J., Sarkar-Dutta, M., Lane, S., Nagaraj, A., Ali, N., et al. (2015). Hepatitis C virus-induced NLRP3-inflammasome activates the sterol regulatory element binding protein (SREBP) and regulates lipid metabolism. *J. Biol. Chem.* 291, 3254–3267. doi: 10.1074/jbc.M115.694059

Memon, R. A., Tecott, L. H., Nonogaki, K., Beigneux, A., Moser, A. H., Grunfeld, C., et al. (2000). Up-regulation of peroxisome proliferator-activated receptors (PPAR-alpha) and PPAR-gamma messenger ribonucleic acid expression in the liver in murine obesity: troglitazone induces expression of PPAR-gamma-responsive adipose tissue-specific genes in the liver of obese diabetic mice. *Endocrinology* 141, 4021–4031.

Meredith, L. W., Wilson, G. K., Fletcher, N. F., and Mckeating, J. A. (2012). Hepatitis C virus entry: beyond receptors. *Rev. Med. Virol.* 22, 182–193. doi: 10.1002/rmv.723

Miyanari, Y., Atsuzawa, K., Usuda, N., Watashi, K., Hishiki, T., Zayas, M., et al. (2007). The lipid droplet is an important organelle for hepatitis C virus production. *Nat. Cell Biol.* 9, 1089–1097. doi: 10.1038/ncb1631

Moreira, S., Correia, M., Soares, P., and Maximo, V. (2011). GRIM-19 function in cancer development. *Mitochondrion* 11, 693–699. doi: 10.1016/j.mito.2011.05.011

Nallar, S. C., Kalakonda, S., Sun, P., and Kalvakolanu, D. V. (2008). GRIM-19: a double-edged sword that regulates anti-tumor and innate immune responses. *Transl. Oncogenomics* 3, 67–79.

Negro, F., and Sanyal, A. J. (2009). Hepatitis C virus, steatosis and lipid abnormalities: clinical and pathogenic data. *Liver Int.* 29(Suppl. 2), 26–37. doi: 10.1111/j.1478-3231.2008.01950.x

Nishimura, T., Kohara, M., Izumi, K., Kasama, Y., Hirata, Y., Huang, Y., et al. (2009). Hepatitis C virus impairs p53 via persistent overexpression of 3beta-hydroxysterol Delta24-reductase. *J. Biol. Chem.* 284, 36442–36452. doi: 10.1074/jbc.M109.043232

Panda, S., Vedagiri, D., Viveka, T. S., and Harshan, K. H. (2014). A unique phosphorylation-dependent eIF4E assembly on 40S ribosomes co-ordinated by hepatitis C virus protein NS5A that activates internal ribosome entry site translation. *Biochem. J.* 462, 291–302. doi: 10.1042/BJ20131530

Park, C. Y., Jun, H. J., Wakita, T., Cheong, J. H., and Hwang, S. B. (2009). Hepatitis C virus nonstructural 4B protein modulates sterol regulatory element-binding protein signaling via the AKT pathway. *J. Biol. Chem.* 284, 9237–9246. doi: 10.1074/jbc.M808773200

Pisonero-Vaquero, S., Garcia-Mediavilla, M. V., Jorquera, F., Majano, P. L., Benet, M., Jover, R., et al. (2014). Modulation of PI3K-LXRalpha-dependent lipogenesis mediated by oxidative/nitrosative stress contributes to inhibition of HCV replication by quercetin. *Lab. Invest.* 94, 262–274. doi: 10.1038/labinvest.2013.156

Reeves, M. B., Davies, A. A., Mcsharry, B. P., Wilkinson, G. W., and Sinclair, J. H. (2007). Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science* 316, 1345–1348. doi: 10.1126/science.1142984

Ruedo-Rincon, N., Bloch, K., Derua, R., Vyas, R., Harms, A., Hankemeier, T., et al. (2015). p53 attenuates AKT signaling by modulating membrane phospholipid composition. *Oncotarget* 6, 21240–21254. doi: 10.18632/oncotarget.4067

Sanders, F. W., and Griffin, J. L. (2016). De novo lipogenesis in the liver in health and disease: more than just a shunting yard for glucose. *Biol. Rev. Camb. Philos. Soc.* 91, 452–468. doi: 10.1111/brv.12178

Seo, T., Lee, D., Shim, Y. S., Angell, J. E., Chidambaram, N. V., Kalvakolanu, D. V., et al. (2002). Viral interferon regulatory factor 1 of Kaposi's sarcoma-associated herpesvirus interacts with a cell death regulator, GRIM19, and inhibits interferon/retinoic acid-induced cell death. *J. Virol.* 76, 8797–8807. doi: 10.1128/JVI.76.17.8797-8807.2002

Shulga, N., and Pastorino, J. G. (2012). GRIM-19-mediated translocation of STAT3 to mitochondria is necessary for TNF-induced necroptosis. *J. Cell Sci.* 125, 2995–3003. doi: 10.1242/jcs.103093

Su, A. I., Pezacki, J. P., Wodicka, L., Brideau, A. D., Supekova, L., Thimme, R., et al. (2002). Genomic analysis of the host response to hepatitis C virus infection. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15669–15674. doi: 10.1073/pnas.202608199

Sun, J. M., Kim, S. J., Kim, G. W., Rhee, J. K., Kim, N. D., Jung, H., et al. (2012). Inhibition of hepatitis C virus replication by Monascus pigment derivatives that interfere with viral RNA polymerase activity and the mevalonate biosynthesis pathway. *J. Antimicrob. Chemother.* 67, 49–58. doi: 10.1093/jac/dkr432

Sung, P. S., Cheon, H., Cho, C. H., Hong, S. H., Park, D. Y., Seo, H. I., et al. (2015). Roles of unphosphorylated ISGF3 in HCV infection and interferon responsiveness. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10443–10448. doi: 10.1073/pnas.1513341112

Suzuki, T. (2012). Morphogenesis of infectious hepatitis C virus particles. *Front. Microbiol.* 3:38. doi: 10.3389/fmicb.2012.00038

Syed, G. H., Amako, Y., and Siddiqui, A. (2010). Hepatitis C virus hijacks host lipid metabolism. *Trends Endocrinol. Metab.* 21, 33–40. doi: 10.1016/j.tem.2009.07.005

Vallianou, I., Dafou, D., Vassilaki, N., Mavromara, P., and Hadzopoulou-Cladaras, M. (2016). Hepatitis C virus suppresses hepatocyte nuclear factor 4 alpha, a key regulator of hepatocellular carcinoma. *Int. J. Biochem. Cell Biol.* 78, 315–326. doi: 10.1016/j.biocel.2016.07.027

Wakita, T., Pietschmann, T., Kato, T., Date, T., Miyamoto, M., Zhao, Z., et al. (2005). Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.* 11, 791–796. doi: 10.1038/nm1268

Waris, G., Felmlee, D. J., Negro, F., and Siddiqui, A. (2007). Hepatitis C virus induces proteolytic cleavage of sterol regulatory element binding proteins and stimulates their phosphorylation via oxidative stress. *J. Virol.* 81, 8122–8130. doi: 10.1128/JVI.00125-07

Xiang, Z., Qiao, L., Zhou, Y., Babiuk, L. A., and Liu, Q. (2010). Hepatitis C virus nonstructural protein-5A activates sterol regulatory element-binding protein-1c through transcription factor Sp1. *Biochem. Biophys. Res. Commun.* 402, 549–553. doi: 10.1016/j.bbrc.2010.10.081

Xu, G., Yang, F., Ding, C. L., Wang, J., Zhao, P., Wang, W., et al. (2014). MiR-221 accentuates IFNs anti-HCV effect by downregulating SOCS1 and SOCS3. *Virology* 46, 343–350. doi: 10.1016/j.virol.2014.06.024

Yang, W., Hood, B. L., Chadwick, S. L., Liu, S., Watkins, S. C., Luo, G., et al. (2008). Fatty acid synthase is up-regulated during hepatitis C virus infection and regulates hepatitis C virus entry and production. *Hepatology* 48, 1396–1403. doi: 10.1002/hep.22508

Yoshida, T., Hanada, T., Tokuhisa, T., Kosai, K., Sata, M., Kohara, M., et al. (2002). Activation of STAT3 by the hepatitis C virus core protein leads to cellular transformation. *J. Exp. Med.* 196, 641–653. doi: 10.1084/jem.20012127

Zhang, W., Niu, M., Yan, K., Zhai, X., Zhou, Q., Zhang, L., et al. (2013). Stat3 pathway correlates with the roles of leptin in mouse liver fibrosis and sterol regulatory element binding protein-1c expression of rat hepatic stellate cells. *Int. J. Biochem. Cell Biol.* 45, 736–744. doi: 10.1016/j.biocel.2012.12.019

Zhao, L. J., He, S. F., Wang, W., Ren, H., and Qi, Z. T. (2016). Interferon alpha antagonizes STAT3 and SOCS3 signaling triggered by hepatitis C virus. *Cytokine* 80, 48–55. doi: 10.1016/j.cyto.2015.08.264

Zhou, Y., Wei, Y., Zhu, J., Wang, Q., Bao, L., Ma, Y., et al. (2011). GRIM-19 disrupts E6/E6AP complex to rescue p53 and induce apoptosis in cervical cancers. *PLoS ONE* 6:e22065. doi: 10.1371/journal.pone.0022065

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kim, Sung, Lee, Hur, Park, Shin, Windisch and Yoon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

REVIEW
published: 19 April 2016
doi: 10.3389/fmicb.2016.00496

Zika Virus: the Latest Newcomer

Juan-Carlos Saiz, Ángela Vázquez-Calvo, Ana B. Blázquez, Teresa Merino-Ramos, Estela Escribano-Romero and Miguel A. Martín-Acebes*

Department of Biotechnology, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

Since the beginning of this century, humanity has been facing a new emerging, or re-emerging, virus threat almost every year: West Nile, Influenza A, avian flu, dengue, Chikungunya, SARS, MERS, Ebola, and now Zika, the latest newcomer. Zika virus (ZIKV), a flavivirus transmitted by *Aedes* mosquitoes, was identified in 1947 in a sentinel monkey in Uganda, and later on in humans in Nigeria. The virus was mainly confined to the African continent until it was detected in south-east Asia the 1980's, then in the Micronesia in 2007 and, more recently in the Americas in 2014, where it has displayed an explosive spread, as advised by the World Health Organization, which resulted in the infection of hundreds of thousands of people. ZIKV infection was characterized by causing a mild disease presented with fever, headache, rash, arthralgia, and conjunctivitis, with exceptional reports of an association with Guillain–Barre syndrome (GBS) and microcephaly. However, since the end of 2015, an increase in the number of GBS associated cases and an astonishing number of microcephaly in fetus and new-borns in Brazil have been related to ZIKV infection, raising serious worldwide public health concerns. Clarifying such worrisome relationships is, thus, a current unavoidable goal. Here, we extensively review what is currently known about ZIKV, from molecular biology, transmission routes, ecology, and epidemiology, to clinical manifestations, pathogenesis, diagnosis, prophylaxis, and public health.

Keywords: Zika, flavivirus, outbreak, microcephaly, zoonosis

OPEN ACCESS

Edited by:
Abraham L. Brass, University of Massachusetts Medical School, USA

Reviewed by:
Sarah Rowland-Jones, University of Oxford, UK
Manoj N. Krishnan, Duke-Nus Graduate Medical School, Singapore

****Correspondence:***
Juan-Carlos Saiz
jcsaiz@inia.es

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *23 February 2016*
Accepted: *27 March 2016*
Published: *19 April 2016*

Citation:
Saiz J-C, Vázquez-Calvo Á, Blázquez AB, Merino-Ramos T, Escribano-Romero E and Martín-Acebes MA (2016) Zika Virus: the Latest Newcomer. Front. Microbiol. 7:496. doi: 10.3389/fmicb.2016.00496

THE VIRUS

Zika virus (ZIKV) is an arbovirus (arthropod-borne virus) classified into the *Flavivirus* genus within the *Flaviviridae* family[1]. Flaviviruses are small enveloped single stranded positive RNA viruses that include important human and animal pathogens such as yellow fever virus (YFV), dengue virus (DENV), West Nile virus (WNV), St. Louis encephalitis virus (SLEV), Japanese encephalitis virus (JEV) or tick-borne encephalitis virus (TBEV) (Gould and Solomon, 2008). Historically, ZIKV was discovered in the course of investigations designed to study the vector responsible for the non-human cycle of yellow fever in Uganda almost 70 years ago. The first isolation was made in April 1947 from the serum of a febrile sentinel rhesus monkey (named Rhesus 766) that was caged in the canopy of Zika Forest, near Lake Victoria (Dick et al., 1952). The second isolation was made from *Aedes africanus* mosquitoes caught in the same forest in January 1948 (Dick et al., 1952). Thus, ZIKV received its name from the geographical area where the initial isolations were made. Both isolations were performed by intracerebral inoculation into *Swiss* albino mice of the samples containing the virus (serum from febrile monkey or mosquito

[1]http://www.ictvonline.org/virustaxonomy.asp

homogenates) demonstrating that ZIKV was a filterable transmissible agent (Dick et al., 1952). These early filtration studies indicated that the size of ZIKV was in the range of about 30–45 nm in diameter (Dick, 1952). Further transmission electron microscopy analysis of ZIKV infected cells revealed that the virions were spherical particles with an overall diameter of 40–43 nm and a central electron dense core being 28–30 nm in diameter (Bell et al., 1971; Hamel et al., 2015). Although there are still no specific studies on the structure of ZIKV, it can be inferred from other flaviviruses (Mukhopadhyay et al., 2005) that the viral particles should be about 50 nm in diameter, which is compatible with the observations performed for ZIKV. Cryoelectron microscopy reconstructions of flavivirus particles have shown that virions are composed by a central core that contains the capsid or core (C) protein associated with the viral genomic RNA. This nucleocapsid is enclosed into a lipid bilayer derived from the host cell. The membrane (M) and envelope (E) proteins are anchored into the lipid envelope and conform the smooth outer shell of the virion, which is constituted by 180 copies of the M and E proteins arranged as 90 anti-parallel homodimers (Kuhn et al., 2002; Mukhopadhyay et al., 2003). Regarding the stability of the virion, it has been described that ZIKV suspensions were most stable at pH of 6.8–7.4 and particles were inactivated at pH of under 6.2 and over 7.8, by potassium permanganate, ether, and temperatures of 58 °C for 30 min, or 60°C for 15 min, but the infectivity was not effectively neutralized with 10% ethanol (Dick, 1952).

Genome

The flavivirus genome is constituted by a single-stranded RNA molecule of positive polarity that, in a similar manner to cellular mRNAs, includes a cap structure at its 5′ end (Dong et al., 2014). Proper methylation of this structure is important not only for efficient translation of viral genome, but also for evasion of immune response (Daffis et al., 2010). The sequence of the prototype strain of ZIKV MR766, which corresponds to a passaged virus derived from the initial ZIKV isolated by intracerebral inoculation of the serum of the febrile monkey (Rhesus 766) into mice in 1947 (Dick, 1952; Dick et al., 1952), revealed that the ZIKV genome was 10794 nucleotides in length (Kuno and Chang, 2007). The genome contains a single open reading frame (ORF) that encodes a polyprotein of about 3400 amino acids (**Figure 1**) that is expected to be cleaved into the mature viral proteins (see next section for polyprotein processing). The single ORF is flanked by two untraslated regions (UTR) located at the 5′ and 3′ ends of the genome, which in the prototype ZIKV MR766 are of 106 and 428 nucleotides in length, respectively (Kuno and Chang, 2007). Remarkably, and in contrast to cellular mRNAs, ZIKV genome lacks a 3′ poly(A) tract and ends with CU_{OH} in a similar manner to the other flaviviruses. Subsequent studies have confirmed that this basic organization is shared among other isolates of ZIKV, although differences in length and nucleotide sequence have been documented among different isolates, even among ZIKV MR766 isolates with different passage history (Lanciotti et al., 2008; Haddow et al., 2012; Baronti et al., 2014; Berthet et al., 2014). The cyclization of flavivirus genome between 5′ and 3′ terminal regions, which is important for the functionality of the genome, is mediated by the interaction of complementary sequences located with genome regions termed conserved sequences (CSs). These CS (CS1 to CS3) are also present in the ZIKV genome, suggesting that has the potential for cyclization. Nevertheless, it has to be remarked that the organization of the CS in the 3′end of ZIKV is different from that of other mosquito-borne flaviviruses (Kuno and Chang, 2007).

Viral Proteins

The viral polyprotein encoded by the single ORF in ZIKV (**Figure 1**), as in other related flaviviruses, is supposed to be cleaved by cellular and viral proteases into three structural proteins: the capsid (C), premembrane/membrane (prM/M), and envelope (E), and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). The predicted cleavage sites of ZIKV basically follow the patterns established for other mosquito-borne viruses, and, likewise, cysteine residues within the polyprotein are well conserved relative to other mosquito-borne flavivirus (Kuno and Chang, 2007). Different proteases participate in the processing of the viral polyprotein: the host cellular signalase which cleaves M/E, E/NS1, and the C-terminal hydrophobic region of NS4A (termed 2K peptide)/NS4B. The viral serine protease (NS3) is expected to cleave the junctions between the virion capsid protein (C_v) and the C-terminal hydrophobic domain of capsid protein (C_i) [C_v/C_i], NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/2K peptide, and NS4B/NS5. The NS1/NS2A is believed to be cleaved by an unknown cellular signalase (Kuno and Chang, 2007). Of key importance is the proteolytic cleavage of prM to give the pr peptide and M protein, which is produced by furin-like protease located in the *trans*-Golgi network during the egress of the particles and promote the maturation of the virions (Mukhopadhyay et al., 2005).

The analysis of the polyprotein sequence predicts the presence of potential N-glycosylation sites in the ZIKV proteins prM, E, and NS1 (Kuno and Chang, 2007; Baronti et al., 2014; Berthet et al., 2014). However, the functional significance of the *N*-glycosylations is not clear in related flaviviruses, since deglycosylated flaviviruses can maintain the same antigenicity, suggesting that carbohydrate does not play a major role in the antigenic properties of the virus (Winkler et al., 1987), and that glycosylation does not alter epitope recognition (Vorndam et al., 1993). On the other hand, glycosylation could be important for replication and maturation (Li et al., 2006). In the case of ZIKV, there are differences among strains due to a 12 nucleotides deletion on the glycosylation motif located at position 154 in the E protein (E-154), which is present in many flaviviruses (Lanciotti et al., 2008; Haddow et al., 2012; Baronti et al., 2014). Remarkably, there are differences on this site even between ZIKV isolates with different passage history, such as those of the prototypic strain ZIKV MR766, indicating that passage history influences glycosylation sites (Haddow et al., 2012). However, the loss of glycosylation on the E-154 residue is not unique to ZIKV and has also been observed in other flaviviruses (Adams et al., 1995; Berthet et al., 1997). Although the functional role of glycosylation in the E protein is not clear, the presence of this glycosylation in

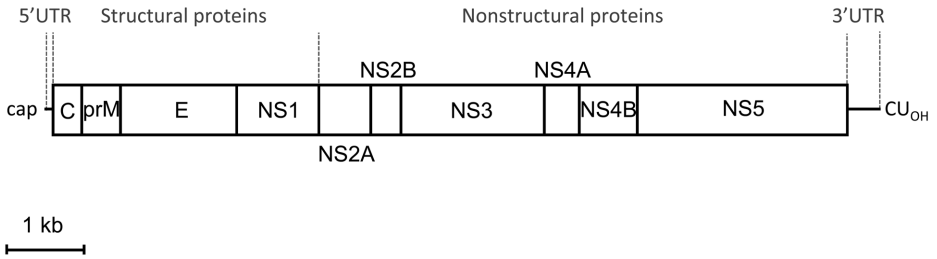


FIGURE 1 | Schematic view of Zika virus (ZIKV) genome organization. The open reading frame (ORF) (boxes) that encodes structural and non-structural proteins is flanked by two untranslated regions (UTR). The proportion of each region was calculated from the ZIKV MR766 sequence available at GenBank (NC_012532.1). Scale bar: 1 kb.

other flaviviruses has been associated with the ability to cause significant human outbreaks (Shirato et al., 2004). Moreover, it has been suggested that the fact that ZIKV strains isolated during the recent human outbreak in Oceania contain this N-linked glycosylation signal, whereas the majority of other strains does not, could indicate that the N-linked glycosylation of the E protein plays a role on the pathogenicity of ZIKV (Baronti et al., 2014; Berthet et al., 2014). Nevertheless, functional studies are required to provide an experimental confirmation of this hypothesis.

Regarding the functions of the flaviviral proteins, the three structural proteins participate in the assembly of the virions. As commented above, the C protein associates with the genomic RNA to conform the core of the virions, and the E protein should mediate the binding to the cellular receptor of the virus and promotes the fusion of the virions with the endosomal membranes of the target cell during viral entry (Mukhopadhyay et al., 2005; Roby et al., 2015). Relative to prM function, this protein assists the folding of E protein as a sort of chaperone and prevents premature fusion of the particles prior to be released from the infected cell, and the cleavage of prM into M protein also promotes the maturation of the viral particles (Mukhopadhyay et al., 2005; Roby et al., 2015).

To our knowledge there are currently no specific studies addressing the function of non-structural proteins of ZIKV, but it is expected that some functions could be inferred from related flaviviruses (Martin-Acebes and Saiz, 2012; Acosta et al., 2014), such as the induction of membrane rearrangements associated with flavivirus replication (NS4A), and the immunomodulation (NS1, NS2A) or regulation of RNA replication and viral assembly (NS2A). Furthermore, NS2B acts as a cofactor for the viral trypsin-like serine protease NS3, which can also act as a helicase. NS5 is the viral RNA dependent RNA polymerase that is in charge of genome replication, and that also displays a methyltransferase domain necessary for capping the 5′ end of the viral genomic RNA. Since flavivirus non-structural proteins constitute major targets for antiviral research (Noble and Shi, 2012; Luo et al., 2015), deciphering the specific role of the non-structural proteins in ZIKV infection could greatly contribute to the development of antiviral strategies against this pathogen.

Host Cell-Virus Interactions

Zika virus can infect a broad range of cells from different tissues and species. For instance, experimental infection by blood meal has revealed that ZIKV replicates in the midgut and salivary glands of diverse *Aedes* mosquitoes (Li et al., 2012; Wong et al., 2013), and also *in vitro* in cultured mosquito cells C6/36 (Hamel et al., 2015). ZIKV also replicates in a wide variety of mammalian cell types. Experimental infection in mice has revealed that the virus replicates mainly in brain cells, including neurons, and astroglial cells (Weinbren and Williams, 1958; Bell et al., 1971), and, *in vitro,* it can replicate in cultured monkey cell lines such as LLC- MK2, or Vero, inducing cytopathic effect (Way et al., 1976). Furthermore, the titer of ZIKV in cultured cells seems to well correlate with the infectivity of the virus *in vivo* (Way et al., 1976). In addition, it has also been recently reported that ZIKV can replicate in human skin cells and also in immature dendritic cells (Hamel et al., 2015). This ability of the virus to replicate in cells from different sources could be related to its transmission cycle, which includes replication in mosquito (vector) and mammalian cells (host).

It has been described that ZIKV enters the cell using adhesion factors such as DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) and diverse members of the phosphatidylserine receptor family (Hamel et al., 2015). Once the attached viral particles are internalized into the cell (**Figure 2**), the viral genome should be released inside the cytoplasm to start translation and replication. The mechanism of penetration of the flavivirus genome into the cytoplasm is initiated by the fusion of the viral envelope with the membranes of the cellular endosomes from the host cell, a process triggered by acidic pH inside cellular endosomes (Stiasny et al., 2011; Vazquez-Calvo et al., 2012). This mechanism of penetration is consistent with, as mentioned before, an early observation showing that ZIKV particles were sensible to acidic pH, and were inactivated by treatment with acidic pH lower than 6.2 (Dick, 1952). Along this line, the sensitivity of ZIKV particles to acidic pH is consistent with the observation performed with other flaviviruses indicating that, in the absence of target membranes, the exposure of flavivirus virions to acidic pH induces rearrangements of the E glycoprotein, which result in a loss of infectivity (Gollins and Porterfield, 1986). The viral RNA acts as mRNA inside the cytoplasm of the

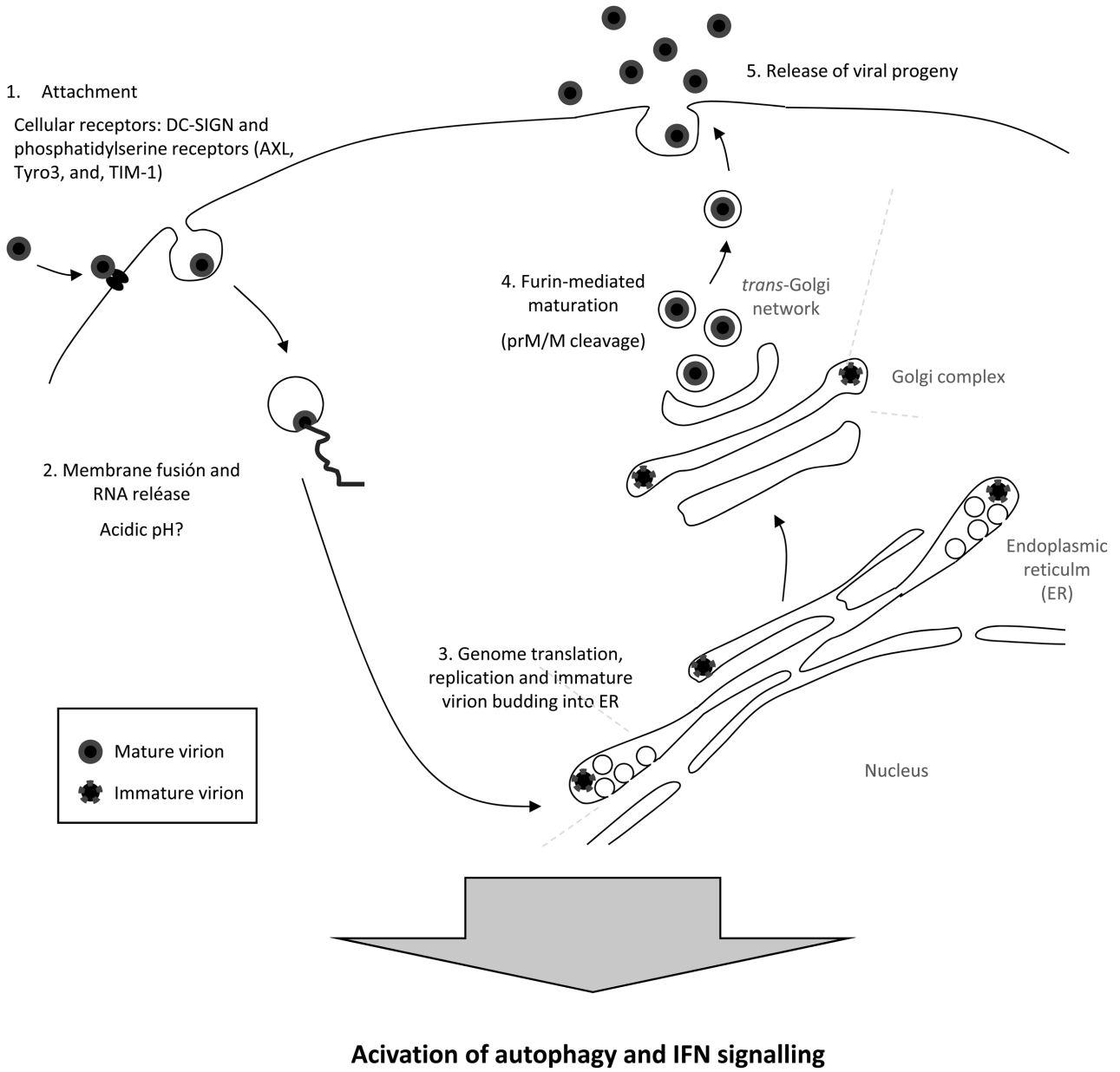


FIGURE 2 | Schematic view of the ZIKV cellular lifecycle. See text for details.

infected cell, and negative-strand viral RNA is synthesized and directs positive-strand RNA synthesis in association with a virus-induced network of membranes derived from the endoplasmic reticulum, ER (**Figure 2**). Electron microscopy studies of virus infected cells showed that ZIKV virions are found in short chains within tubular elements of the ER, which appeared to be in continuity with distended cisternae (Bell et al., 1971). These images were similar to membrane rearrangements observed in other flavivirus infected cells (Welsch et al., 2009; Martin-Acebes et al., 2011; Miorin et al., 2013). Although flavivirus replication is thought to occur in the cellular cytoplasm, it should be noted that one study reported that ZIKV antigens could be found in infected cell nuclei (Buckley and Gould, 1988).

De novo synthesized positive strand-RNA has to be packaged in progeny virions that bud into the ER to form enveloped immature virions (**Figure 2**). These virions traffic through the Golgi complex and, then, the prM is cleaved in the *trans*-Golgi network for particle maturation prior to release from the infected cell (Mukhopadhyay et al., 2005; Roby et al., 2015). Remarkably, the observations made for other flaviviruses have revealed that not all the copies of the prM protein in the secreted virions are cleaved, and, thus, that a proportion of them remains unprocessed in the secreted virions (Plevka et al., 2011). Even

more, the amount of prM within the flavivirus virions varies with the cell line used for the production of the virus and the infecting flavivirus, and could be of key importance for the antigenicity of the particles (Pierson and Diamond, 2012; Lok, 2016).

Knowledge regarding the cellular response to ZIKV infection is still scarce. However, it has been experimentally probed that replication of ZIKV provokes an innate antiviral response, inducing the transcription of TLR3, RIG-I, and MDA5, as well as that of several interferon stimulated genes, such as OAS2, ISG15, and MX1, characterized by a strongly enhanced beta interferon gene expression (Hamel et al., 2015). In addition, ZIKV infection is sensitive to interferon (IFN) signaling, as pretreatment of primary skin fibroblasts with IFN-alpha, beta, and gamma reduces ZIKV infection (Hamel et al., 2015). ZIKV infection also upregulates the autophagic pathway in infected skin fibroblasts (Hamel et al., 2015), which is consistent with the observations for other related flaviviruses (Blazquez et al., 2014). Moreover, the autophagic marker LC3 colocalizes with viral proteins within ZIKV-infected cells, and the infection can be reduced by treatment with the autophagy inhibitor 3-methyladenine, whereas upregulation of autophagy using Torin 1 increases ZIKV replication (Hamel et al., 2015).

MOLECULAR CLASSIFICATION

Zika virus is genetically and antigenically related to Spondweni virus. Both viruses form a unique clade (clade X) within the mosquito-borne flavivirus cluster (Kuno et al., 1998) (**Figure 3**). Phylogenetic analyses reveal the existence of two major lineages: one includes the African strains, and the other the Asian and American strains (Haddow et al., 2012; Alera et al., 2015) (**Figure 4**). The African lineage is further divided into two groups, the East African cluster, containing the genetic variants of the

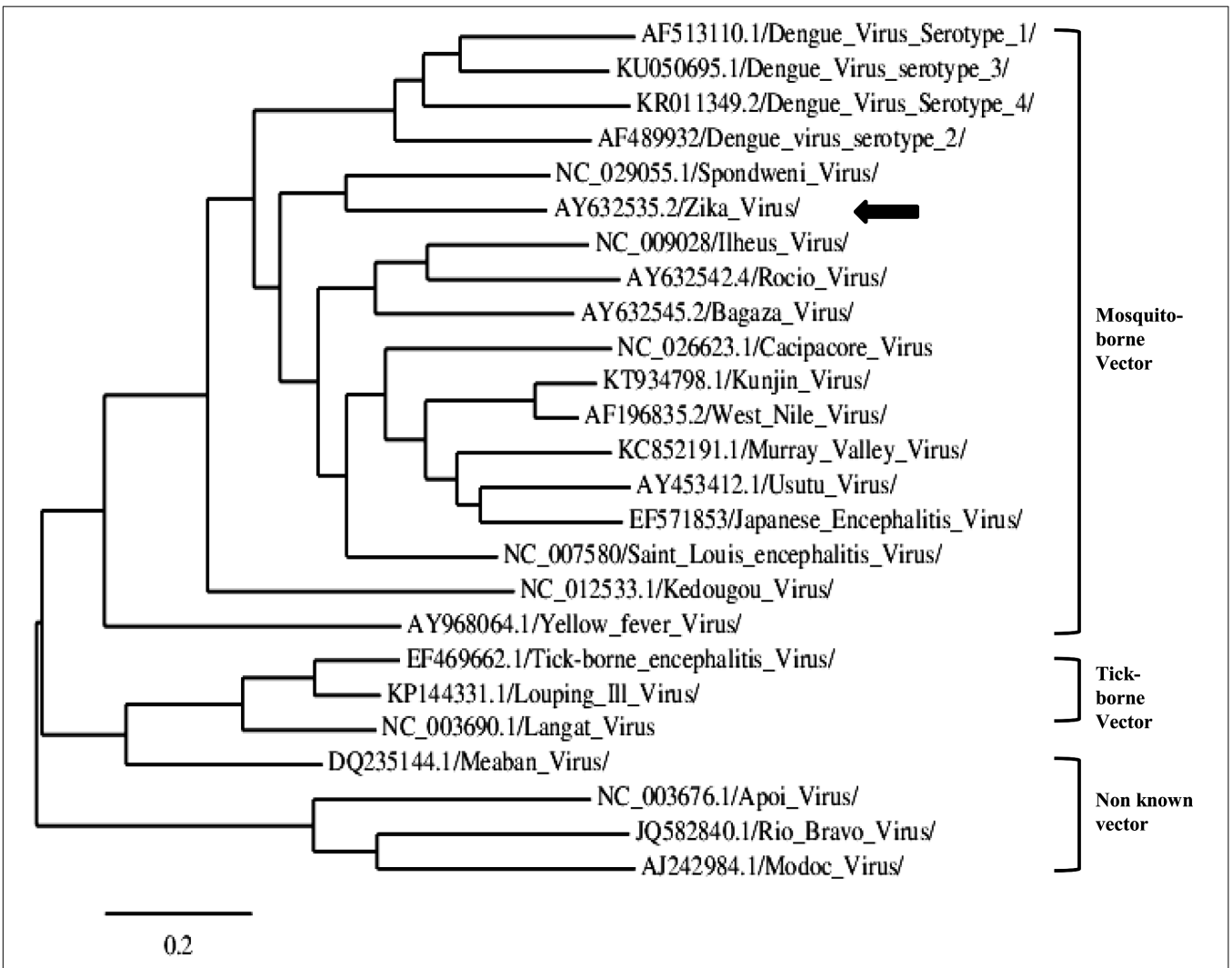


FIGURE 3 | Representative phylogram showing the relationships between strains of the genus Flavivirus. Accession numbers are displayed in the tree. The arrow indicates ZIKV. The scale indicates 0.2 substitutions/site. The tree was based on complete NS5 nucleotide sequence, built from a multiple alignment using Clustal omega and Phylogeny.fr (Dereeper et al., 2008).

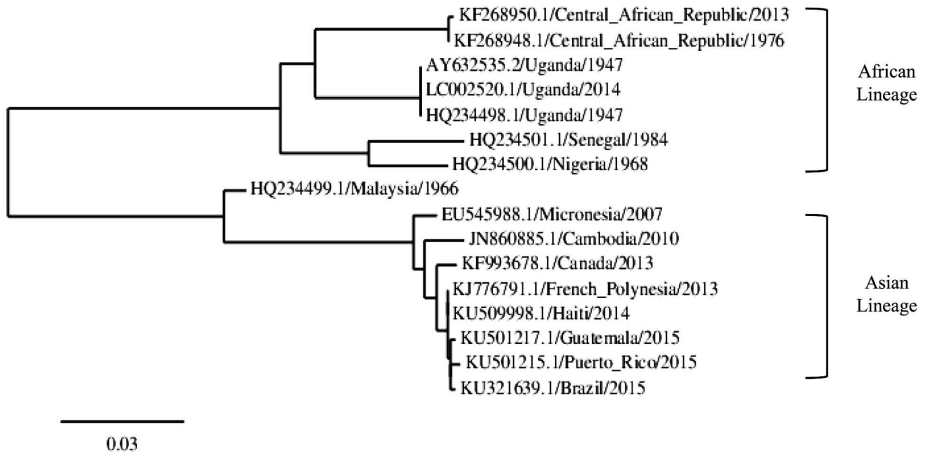


FIGURE 4 | Representative phylogram showing the relationships between ZIKV strains. The accession number, country of isolation and year is displayed in the tree. The scale indicates 0.03 substitutions/site. The tree was based on complete NS5 nucleotide sequence, built from a multiple alignment using Clustal omega and Phylogeny.fr.

prototypic MR766 strain isolated in Uganda in 1947, and a second group including West African strains (Olson et al., 1981).

Phylogenetic studies (Faye et al., 2014) have established the date of the emergence of ZIKV in east Africa around 1920 (confidence range of 1892–1947). The same study dated the transmission of eastern African ZIKV to Asia around 1945 (confidence range 1920–1960), where the virus was first detected in the late 1960s in Malaysia (Marchette et al., 1969), and subsequently across south-east Asia. These data indicated a widespread occurrence of ZIKV from Africa to Southeast Asia, west and north of the Wallace line (Lanciotti et al., 2008).

Phylogenetic studies have confirmed that Pacific Island ZIKV strains are related to the Asian lineages (Gatherer and Kohl, 2016). Due to the great geographical distances involved, it seems likely that the virus was introduced to the island either by a viremic person, an enzootic host species, or an infected mosquito transported to the island (Haddow et al., 2012). On the other hand, recent transmission to the Americas appears to have originated in the Pacific Islands (Campos et al., 2015; Zanluca et al., 2015). Phylogenetic analysis of the sequences placed the Brazilian and other American strains in a clade with sequences from the Asian lineage, showing a 99% identity with a sequence from a ZIKV isolate from French Polynesia (Baronti et al., 2014; Campos et al., 2015) (**Figure 4**). It has been postulated that two events may have led to the introduction of ZIKV in Brazil, the 2014 FIFA World Cup tournament and an international canoe racing event (Musso, 2015), but since Pacific nations were only represented among the canoe racers, the latter seems to be a likeliest introduction route.

As exemplified in **Figure 4**, ZIKV strains collected in the same geographical region during several years show minimal changes on their sequences, as is the case of strains collected from mosquitoes in a 3 years interval in Central African Republic (Berthet et al., 2014). In this regard, it has been described that infection and transmission modes of ZIKV allow the accumulation of synonymous mutations and negatively selected certain sites (Faye et al., 2014). The arbovirus life cycle imposes several barriers to non-synonymous mutations in some important genes as a consequence of the intrinsic constraints associated with dual replication in mammalian and invertebrate hosts, thus driving to a more slowly fixation of mutations of these viruses when compared with RNA viruses transmitted by other routes. In fact, arboviruses, which are able to successfully adapt to diverse cell types, are characterized by a high rate of deleterious mutations (Holmes, 2003).

Human-to-human transmission of the Asian ZIKV strains along the Pacific Islands and South America has been associated with significant NS1 codon usage adaptation to human housekeeping genes, which could facilitate viral replication and increase viral titers (Freire et al., 2015). Furthermore, this report predicted the presence of several epitopes in the NS1 protein that are shared between ZIKV and DENV, pointing to a significant dependence of the recent human ZIKV spread on NS1 translational selection.

As mentioned above, it is noteworthy to note that several of the ZIKV strains exhibited a 4 amino acid deletion corresponding to the envelope protein 154 glycosylation motif found in many flaviviruses (Berthet et al., 2014).

TRANSMISSION CYCLE

Mosquitoes

The arthropod vectors of the ZIKV natural transmission cycle are mosquitoes of the genus *Aedes* (Diagne et al., 2015). As mentioned early, the virus was first isolated from *A. africanus* (Dick et al., 1952) and, since then, ZIKV has been isolated from *A. aegypti* (Marchette et al., 1969), the main vector of the virus, and also from *A. albopictus* (Grard et al., 2014), confirming that both species are competent vectors.

Aedes aegypti is currently distributed in Asia and Oceania, the Americas, and in a few regions of Africa and Europe (Madeira

and the north-eastern Black Sea coast)[2]; however, it has been recently predicted that this species might soon colonize some southern European regions, as well as temperate North America and Australia (Kraemer et al., 2015). The first evidence of the role of *A. aegypti* in the urban transmission cycle of ZIKV was suggested after its isolation from a pool of mosquitoes collected in 1966 in Malaysia, in what it was the first isolation from a mosquito other than *A. africanus* (Marchette et al., 1969).

Aedes albopictus, the so-called Asian tiger mosquito, is also widely distributed. This species is currently circulating in Asia, North, Central and South America (Kraemer et al., 2015), northern Australia, and in some areas of Africa and southern Europe, where it has spread in the past two decades to France, Germany, Italy, and Spain (Paupy et al., 2009; Dyer, 2016). Contrary to *A. aegypti*, *A. albopictus* can hibernate and survive in cool temperature regions (Thomas et al., 2012). This species can also efficiently transmit the virus, as demonstrated during the outbreak that took place in Gabon in 2007, where, among all species tested (including *A. aegypti*), it was the only one in which the virus was detected, thus confirming that *A. albopictus* may also play an important role in ZIKV transmission (Grard et al., 2014).

The ability of ZIKV to be efficiently transmitted by both mosquito species (*A. aegypti* and *A. albopictus*) that feed on humans further complicates their control and, thus, that of ZIKV. Both species grow very close to human populations, but while *A. aegypti* feed almost exclusively on humans in daylight hours and typically rest indoors (Scott and Takken, 2012), *A. albopictus* is usually exophagic and bites humans and also domestic and livestock animals (Paupy et al., 2009), although under some circumstances it preferentially feed on humans, hence, confirming that it can also have an anthropophilic behavior similar to *A. aegypti* (Ponlawat and Harrington, 2005; Delatte et al., 2010). Therefore, methods of control for a species may not be accurate to control the other one. Furthermore, when the populations of *A. aegypti* is reduced, the opportunistic invasive *A. albopictus* may rapidly move into the area (Higgs, 2016).

Aedes aegypti and, to a lesser extent, *A. albopticus* are clearly involved in ZIKV transmission and spread, but other *Aedes* spp., such as *A. polynesienis*, are suspected to have also contributed to it, as it was the case during the 2013–2014 outbreak in the French Polynesia (Cao-Lormeau et al., 2014). In fact, ZIKV has been also isolated from, at least, other 15 *Aedes* species in different regions of the world (**Table 1**). Even more, in the Kédougou region of Senegal, ZIKV was also amplified from pools of three mosquito species other than *Aedes: Anopheles coustani*, *Culex perfuscus,* and *Mansonia uniformis* (Diallo et al., 2014).

As in other arboviral infections, local overwintering could be an important aspect for maintenance and spread of ZIKV. Detection of the virus in a pool of *A. furcifer* males in 2011 in Senegal, even though no infected *A. furcifer* females were collected, strongly suggested that ZIKV is vertically transmitted, at least in this species, and that this transmission route may be an important mechanism of local maintenance (Diallo et al., 2014).

Finally, although most of these data were obtained by analyses of ZIKV naturally infected mosquitoes, their current vector competence still has to be clearly established. In this sense, different *Aedes* species (*A. aegypti, A. unilineatus, A. vittatus,* and *A. luteocephalus*) were recently tested in their susceptibility to ZIKV oral infection, and, although all of them were susceptible, viral genome could be amplified from saliva only in the case of *A. vittatus* and *A. luteocephalus* mosquitoes (Diagne et al., 2015).

TABLE 1 | Zika vector mosquito species.

Species	Arguments in favor of vector status	Place and year of first report (Reference)
Aedes africanus	Isolation	Zika Forest (Uganda), 1948 (Dick et al., 1952)
Aedes aegypti	Isolation	Malaysia, 1966 (Marchette et al., 1969)
Aedes luteocephalus	Isolation	Saboya Forest (Senegal), 1968 (Haddow et al., 2012; Diallo et al., 2014)
Aedes vittatus	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes dalzieli	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes metallicus	Isolation	Senegal 1968 (Haddow et al., 2012)
Mansonia uniformis	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes fowleri	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes minutus	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes neoafricanus	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes tarsalis	Isolation	Senegal 1968 (Haddow et al., 2012)
Aedes apicoargenteus	Isolation	Central African Republic, 1969 (Diagne et al., 2015)
Aedes furcifer-taylori	Isolation	Bandia (Senegal), 1969 (Diallo et al., 2014)
Annopheles gambiae s.l.	Isolation	Bandia (Senegal), 1969 (Diallo et al., 2014)
Aedes opok	Isolation	Central African Republic, 1976–1980 (Berthet et al., 2014)
Aedes taylori	Isolation	Kédougou region (Senegal), 1988–1991 (Monlun et al., 1993)
Aedes albopictus	Epidemiology	Gabon, 2007 (Grard et al., 2014)
Aedes hirsutus	Epidemiology	Kédougou region (Senegal), 2011, (Diallo et al., 2014)
Aedes unilinaetus	Epidemiology	Kédougou region (Senegal), 2011, (Diallo et al., 2014)
Culex perfuscus	Epidemiology	Kédougou region (Senegal), 2011, (Diallo et al., 2014)
Anopheles coustani	Epidemiology	Kédougou region (Senegal), 2011, (Diallo et al., 2014)

[2]http://ecdc.europa.eu/en/healthtopics/vectors/vector-maps/Pages/VBORNET_maps.aspx

Humans and Non-human Primates

During outbreaks, humans are the primary host for ZIKV (Staples et al., 2016), and both urban (Grard et al., 2014) and sylvatic (Berthet et al., 2014) viral transmission have been demonstrated. First ZIKV isolation from humans was reported in 1954 in Nigeria (Macnamara, 1954), although antibodies to the virus had been previously found during early surveys on human sera in different regions of Africa (Dick, 1952; Smithburn, 1952). Recently, and albeit an estimated 80% of ZIKV infected people are asymptomatic[3], during the ongoing outbreak in Brazil, ZIKV RNA has been identified in brain, placenta, and amniotic fluid specimens, and its presence has been associated to microcephaly in infants and miscarries during pregnancy (Martines et al., 2016; Mlakar et al., 2016).

One important aspect that still has to be clarified is whether ZIKV infection in humans drives to viral titers enough to initiate a new cycle when an infected person is bitten by a naïve mosquito. Early studies attempting to infect *A. aegypti* from a ZIKV-infected human volunteer to further transmit the agent to newborns mice were unsuccessful (Bearcroft, 1956), probably because viremia was too low, even though the volunteer was bitten during the acute phase of the disease (4–6 days after infection). Later on, a recent attempt to isolate circulating virus from infected patients of the 2007 Gabonese outbreak on mammalian (Vero) and insect (C6/36) cell lines was again unsuccessful, presumably because of low viral titers (despite two patients presenting only 1 and 4 days after symptom onset), although the inappropriate initial storage conditions could have also contributed to it (Grard et al., 2014). In this line, the estimated number of genome copies circulating in ZIKV-infected patients during the 2007 outbreak on the Pacific Island of Yap was reported to be 0.9×10^3–7.2×10^5 cDNA copies/ml (Lanciotti et al., 2008). These relatively low levels of viremia among ZIKV-infected patients are very far from those reported for other arboviruses, as Chikungunya (CHIK) and DENV-2, where an estimated 10^7 and 10^8 cDNA copies/ml were reported, respectively (Caron et al., 2012), but they are in the range of other dead-end flavivirus infection in humans, such as that of WNV, where viral loads from 50 to 6.9×10^5 copies/ml are observed (Pupella et al., 2013). Therefore, further experiments are needed to clarify this issue.

In the case of non-human primates, it is known that epizootics occur in them (McCrae and Kirya, 1982), but it is unclear whether they are an obligatory reservoir in the transmission to humans. In Africa, ZIKV natural transmission cycle involves primarily *Cercopithecus aethiops* and *Erythrocebus patas* monkeys (Faye et al., 2013). In Asia (Borneo), antibodies against ZIKV have been detected among semi-captive and wild orangutans (Wolfe et al., 2001). However, this study reported a higher prevalence of anti-ZIKV antibodies in humans than in orangutans, suggesting a possible incidental infection of these animals through contact with mosquitoes infected by viremic people, or from recently established sylvatic cycles. Nonetheless, it is also possible that sylvatic ZIKV transmitting mosquitoes in Borneo have a more narrow distribution, or an ecology that does not lead to frequent exposure by orangutans.

An acute symptomatic ZIKV infection case after a monkey bite has been recently described (Leung et al., 2015). Even though transmission throughout mosquito bite could not be completely discharged, the presence of ZIKV in the human pharynx and the previous identification of ZIKV RNA in saliva from asymptomatic infected patients (Besnard et al., 2014) could be consistent with a potential transmission from primates' bite (Leung et al., 2015).

Other Vertebrates

Information regarding the possible susceptibility of animals other than human and non-human primates is limited. Antibodies directed against ZIKV have been found in several vertebrates species, such rodents, birds, reptiles, goats, sheep, and cattle in Kenya (Johnson et al., 1977), and in Pakistan, where some species of rodents were suggested as possible reservoirs of ZIKV (Darwish et al., 1983). In addition, the rapid periodicity of amplification observed in Senegal along the 2011 outbreak could support that, besides primates, other vertebrates may also play a role in ZIKV circulation (Diallo et al., 2014).

Non-vector

Sporadic reports of direct human-to-human transmission have been reported to occur perinatally, sexually, and through breastfeeding and blood transfusion. Likewise, occupational transmission in the laboratory setting has also been described (Filipe et al., 1973).

Perinatal transmission from two mothers to their new-borns during the French Polynesia outbreak has been documented, although contamination during delivery was not completely discarded (Besnard et al., 2014). Sera from the mothers were positive for ZIKV by reverse transcriptase polymerase chain reaction (RT-PCR) within 2 days post-delivery and those of their newborns within 4 days after birth. In addition, high ZIKV RNA load was detected in breast milk samples from both mothers, albeit the virus could not be multiplied in susceptible cell cultures. Therefore, ZIKV transmission by breastfeeding must be considered and further clarify. Similarly, a first case of perinatal transmission of ZIKV was suspected to have occurred during the same outbreak from a mother that present a ZIKV infection-like syndrome 2 weeks before delivery, with the newborn showing a maculopapular rash at birth; however, virological investigations were not performed in this case (Besnard et al., 2014).

Besides these sporadic cases of non-vector transmission, the most surprising phenomenon on ZIKV infection is, without any doubt, the unexpected number of infants supposedly born with microcephaly (see clinical manifestations and pathogenesis section below) in Brazil during the ongoing viral outbreak, apparently as a result of their mothers being infected during pregnancy, since, in some cases, ZIKV-RNA has been detected in the amniotic fluid of the mothers[4]. Around 4000 cases of Zika-related microcephaly have been recorded in that country until

[3]http://ecdc.europa.eu/en/healthtopics/zika_virus_infection/factsheet-health-professionals/Pages/factsheet_health_professionals.aspx

[4]http://ecdc.europa.eu/en/publications/Publications/rapid-risk-assessment-zika-virus-first-update-jan-2016.pdf

now since October 2015, accounting for over 40 infants deaths (Higgs, 2016). Furthermore, ZIKV has been recently found in the fetal brain tissue of a baby with microcephaly after termination of the pregnancy requested by the infected woman (Mlakar et al., 2016).

At present (February 2016), three reported cases indicate that ZIKV could be sexually transmitted. In August 2008, a scientist was bitten many times while studying mosquitoes in south-eastern Senegal. Six days after returning to his home in Colorado (U.S.), he felt ill with symptoms of Zika fever and hematospermia. By then, he had had unprotected intercourse with his wife, who had not been outside the U.S. during the previous year, which subsequently developed symptoms of Zika fever. Virus infection was confirmed by serologic testing in both, but the presence of ZIKV in the semen of the patient was not investigated (Foy et al., 2011). A second report described the presence of replicative ZIKV and a high ZIKV RNA load in semen ($1{,}1–2{,}9 \times 10^7$ copies/ml) and urine ($3{,}8 \times 10^3$ copies/ml) samples of a patient during the 2013 outbreak in Tahiti, but no RT-PCR amplification was obtained from sera collected at the same time (Musso et al., 2015b). In early February 2016, the Dallas County Health and Human Services Department (U.S.) reported a third case, still under investigation, of a person that apparently contracted Zika fever after sexual contact with an ill person who had recently returned from a ZIKV high risk country[5].

In addition to perinatal, breastfeeding, and sexual sporadic transmission of ZIKV, the potential for viral transmission through blood transfusion was demonstrated during the French Polynesian outbreak. Almost 3% (42/1505) of blood donors, who were asymptomatic at the time of donation, were found positive for acute ZIKV infection by specific RT-PCR (Musso et al., 2014a). Further studies are needed to assess the actual risk of ZIKV transmission through blood products and the risk to generate a disease in the recipient, but these data point to the necessity for quickly adapting blood donation safety procedures to the local epidemiological context. In fact, the Pan American Health Organization, PAHO[6], and the European Centre for Disease Prevention and Controls, ECDC[7], have recently issued a bulletins to alert their national health and blood safety authorities on this still poorly recognized viral infection.

Finally, ZIKV has been also detected in saliva samples (Besnard et al., 2014) with even higher frequency than in blood samples (Musso et al., 2015a) and, thus, saliva is another transmission source that have to be considered.

In resume, although ZIKV transmission by routes other than mosquito bites has been so far sporadic, further studies are mandatory to clearly establish their possible role on ZIKV epidemics.

[5]http://www.dallascounty.org/department/hhs/press/documents/PR2-2-16DCHHSReportsFirstCaseofZikaVirusThroughSexualTransmission.pdf

[6]http://www.paho.org/hq/index.php?option=com_content&view=article&id=11605:2016-paho-statement-on-zika-transmission-prevention-&Itemid=41716&lang=en

[7]http://ecdc.europa.eu/en/press/news/_layouts/forms/News_DispForm.aspx?List=8db7286c-fe2d-476c-9133-18ff4cb1b568&ID=1348

ECOLOGY

Most arboviruses are perpetuated in transmission cycles independent of human hosts, but those with sylvatic cycles often infect people who accidentally intrude on their natural habitats. Nonetheless, in many instances, humans are dead-end hosts in complex transmission cycles that involve different wild and domestic vertebrate hosts, as in the case of JEV (Wolfe et al., 2001) or WNV (Martin-Acebes and Saiz, 2012). Less frequently, arboviruses may jump from this sylvatic transmission cycle to a mainly human-mosquito transmission cycle, as it likely was the case of CHIKV.

In the case of ZIKV, early studies indicated that non-human primates were the primary vertebrate hosts, with occasional participation of humans in the transmission cycle, even in highly enzootic areas. This theory was based on evidences indicating that *A. africanus*, a species with a greater preference for them than for humans (Haddow and Dick, 1948; Haddow et al., 1964), was the principal (if not the only) ZIKV vector (Haddow et al., 1964). However, by now, it seems clear that, as commented before, *A. aegypti* and, to a lesser extent, *A. albopticus* are the main vectors and, thus, that humans probably serve as primary amplification hosts when their viremia is sufficient in duration and magnitude (Haddow et al., 2012).

Although it is still unknown if ZIKV overwinters in geographical areas without reports of cases, it appears that, as in other flaviviral infections, such as that caused by WNV (Martin-Acebes and Saiz, 2012), epidemics are more related to the specific mosquito species, its population density, competence, and behavior in a given area, as this can shape virus dynamics.

Aedes aegypti, which, as mentioned before, is found throughout Asia, Oceania, the Americas and in some regions of Africa and Europe[2] (Kraemer et al., 2015), does not overwinter, but can be sheltered in domestic settings, which provides protection against environmental conditions and numerous aquatic habitats suitable for oviposition. Beside Georgia, this species does not currently circulate in Europe, but there are no climatic reasons to believe that, if re-introduced, it cannot become widely established in southern Europe as before (Reiter, 2010).

Meanwhile, *A. albopictus*, the most invasive mosquito species in the world (Medlock et al., 2012), have spread during the last 30–40 years to North, Central and South America, parts of Africa, southeastern Asia, China, Japan, northern Australia, and southern Europe (Paupy et al., 2009). This successfully colonization of new regions is due to its ability to adapt to different climates through the production of cold-resistant eggs, with temperate strains surviving cold winters in northern latitudes. On top of that, its preference for container habitats (e.g., tires and vases) in domestic settings has resulted in increased potential for contact with humans (Medlock et al., 2012).

ZIKV pandemic is currently in progress, with many important questions still unanswered. However, as seen during recent years with other viral agents, urban overcrowding, constant international travel, disruption of the ecologic balance, and climate changes can favored the unexpected emergence of asleep, or yet unknown, infectious agents (Fauci and Morens, 2016). Therefore, comprehensive and integrated investigations have

to be conducted to better understand the complex ecosystems in which agents of current and future pandemics could be aggressively evolving.

EPIDEMIOLOGY

Africa

Since the first isolation of ZIKV in 1954 from an inhabitant of Nigeria (Macnamara, 1954), many serological and entomological studies have reported the circulation of the virus across a widespread area of Africa, including Kenya (Geser et al., 1970), Nigeria (Lee and Moore, 1972; Monath et al., 1973; Fagbami, 1979), Sierra Leone (Robin and Mouchet, 1975), Gabon (Jan et al., 1978; Grard et al., 2014), Uganda (McCrae and Kirya, 1982), Central African Republic (Saluzzo et al., 1981), Senegal (Monlun et al., 1993; Diallo et al., 2014; Althouse et al., 2015), and Ivory Coast (Akoua-Koffi et al., 2001), with prevalence ranging 1.3–52%. At present, Cape Verde is the only African country where, since last year, active viral transmission is currently being reported[8,9] (**Figure 5**). From the beginning of the outbreak in October 2015 to 7th February of this year, the Health Authorities have reported 7362 cases without associated neurological disorders[10]. In any case, and although until now human cases of ZIKV related disease have only been sporadically documented in Africa, it should be kept in mind that this might have been partially due to underdiagnoses, mainly in areas where DENV and CHIKV circulate, as infection with all these viruses presents similar clinical signs (Weissenbock et al., 2010; Grard et al., 2014).

Asia

In Asia, distinguishing ZIKV infection from other arboviral infections (dengue, yellow fever, and other tropical diseases) is also difficult, so that, early epidemiological data should be treated with caution. In the early 50's, seroprevalence of ZIKV in humans was reported in several countries: India, Malaysia, Philippines, Vietnam, and Thailand with variable rates, 8–75% (Smithburn, 1954; Smithburn et al., 1954; Hammon et al., 1958; Pond, 1963; Marchette et al., 1969). As the time went by, other countries reported seroprevalence in humans: Indonesia (Java island) from 1977 to 1978 (Olson et al., 1981, 1983) and Pakistan in 1980 (Darwish et al., 1983). More recently a few sporadic human cases have been documented: a confirmed case in Cambodia in 2010 (Heang et al., 2012), an infected traveler who returned from Indonesia in 2013 (Kwong et al., 2013), and some several ZIKV positive cases from travelers returning from Thailand during 2012–2014 (Buathong et al., 2015). Very recently, in January 2016, WHO notified a case of ZIKV infection in a traveler coming back to Finland after spending a few months in the Maldives[11] (**Figure 5**).

[8]http://ecdc.europa.eu/en/healthtopics/zika_virus_infection/zika-outbreak/Pages/Zika-countries-with-transmission.aspx

[9]http://www.cdc.gov/zika/geo/active-countries.html

[10]http://www.minsaude.gov.cv/index.php/documentos/cat_view/34-documentacao/72-zika-virus

[11]http://www.who.int/csr/don/8-february-2016-zika-maldives/en/

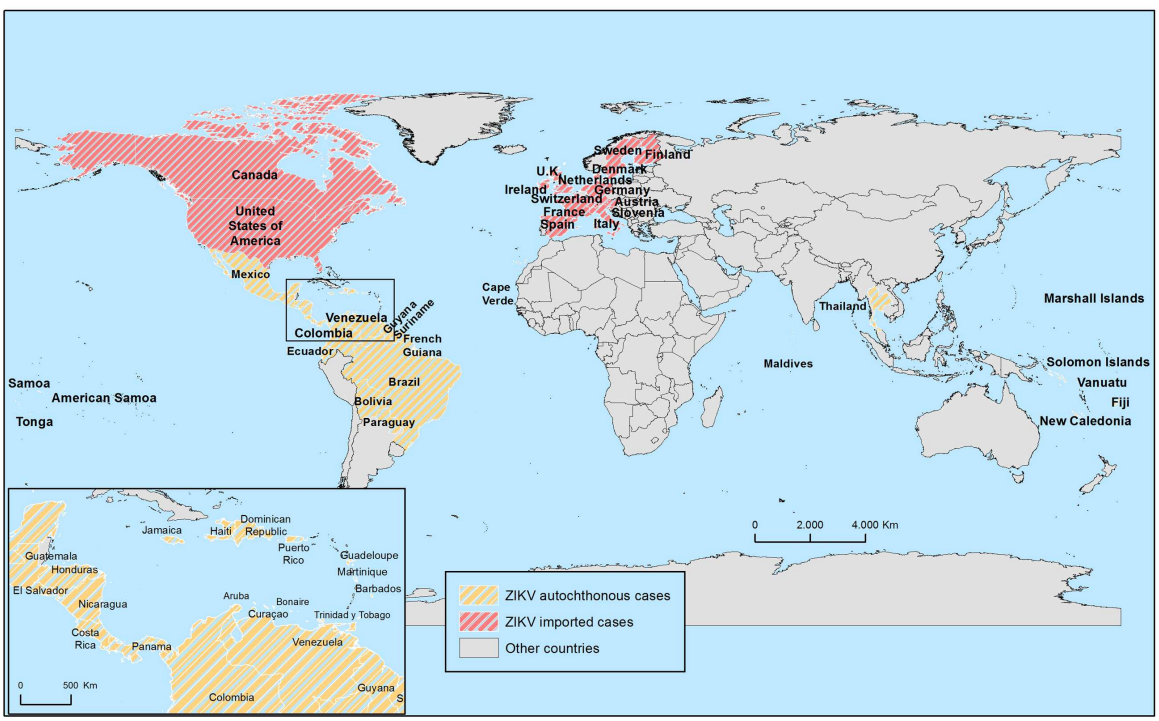


FIGURE 5 | Map showing worldwide autochthonous and imported ZIKV human cases since the last 9 months. See text for details.

Oceania

In April 2007, a ZIKV outbreak was reported in Yap Island, Micronesia. The way through the virus was introduced is still unknown, but, as commented before, it has been proposed it was due to an infected mosquito or to an asymptomatic person with undetected infection. Serological analysis indicated that 73% of Yap residents had been infected with ZIKV, but not a single hospitalization, hemorrhagic manifestation, or death were reported during the outbreak (Duffy et al., 2009).

After detection of the first case of ZIKV infection in the French Polynesia in October 2013, during the following outbreak, up to an estimated 11% of population was affected (Cao-Lormeau et al., 2014; Musso et al., 2014b). Most of the clinical cases presented low fever, asthenia, wrist and fingers arthralgia, headache, rash, and only one patient presented Guillain–Barre syndrome (GBS) 7 days after laboratory confirmation of ZIKV infection. It is noteworthy to remark that, after this outbreak, the incidence of GBS in French Polynesia increased 20-fold (Oehler et al., 2014). The reasons for this increase are not known yet. Analyses of the circulating virus have shown that it was genetically closely related to the 2007 Yap and the 2010 Camboya strains, which were not associated to a noticeable number of GBS cases.

Following the French Polynesia outbreak in late 2013, subsequent outbreaks occurred in New Caledonia, Eastern Island, and the Cook Islands. In New Caledonia, the first cases of ZIKV infection were imported from French Polynesia in November 2013, and, subsequently, in January 2014, the first autochthonous case was documented, driving the New Caledonia Health Authority to declare an outbreak situation in February 2014. Up to 1385 laboratory confirmed cases were reported (Dupont-Rouzeyrol et al., 2015). The outbreak in Eastern Island also started in January 2014, accounting for 51 confirmed cases (Tognarelli et al., 2016).

In Australia, the first case of ZIKV infection was notified in 2012 in a traveler returning from Indonesia (Kwong et al., 2013). Since then, all cases have been imported from countries affected by the virus. Nonetheless, as *A. aegypti*, the main ZIKV vector, is present mainly in areas of North Queensland, this region is at risk of an eventual local transmission of ZIKV from infected returning travelers[12]. A few imported cases have been confirmed in New Zealand too, but as the mosquito vectors are not commonly found in these territories, the risk of local transmission is so far low[13].

In the past 9 months, Fiji, New Caledonia, Vanuatu, Solomon Islands, Marshall Islands, American Samoa, Samoa, and Tonga (**Figure 5**) have reported autochthonous cases[8,9] but only the latter three have ongoing outbreaks[14].

America

At the beginning of 2015, first autochthonous transmission of ZIKV in Brazil was reported and, since then, the virus has rapidly spread throughout the Americas (Zanluca et al., 2015). The Brazilian Ministry of Health estimates that between 440 000 to 1 300 000 cases of ZIKV infections may have occurred in 2015 in the country[15].

Since the initial report in Brazil, during the last 9 months, and as a 18th of February 2016, 29 countries, or territories, of America have also reported autochthonous cases of ZIKV infection: Aruba, Barbados, Bolivia, Bonaire, Brazil, Colombia, Costa Rica, Curaçao, Dominican Republic, Ecuador, El Salvador, French Guyana, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Jamaica, Martinique, Mexico, Nicaragua, Panama, Paraguay, Puerto Rico, Saint Martin, Suriname, Trinidad y Tobago, Venezuela, and Virgin Islands[8] (**Figure 5**). Because the epidemic is still spreading in the Americas, it is reasonable to think that more countries will report autochthonous ZIKV infections in the coming months.

Despite outbreaks have occurred mainly in Caribbean countries/territories and central and south America, imported confirmed cases are starting to show in North America: USA and Canada (**Figure 5**). As of today, 82 and 3 imported cases have been reported in these countries, respectively[16,17]. Although the risk of ZIKV establishment in Canada and northern USA is low because of the absence of the vectors, it cannot be discarded that the virus persists and spreads in these regions causing autochthonous infected human cases, as it has happened with other related flaviviruses like WNV (Martin-Acebes and Saiz, 2012).

Europe

Nowadays, there is no evidence of autochthonous ZIKV infection in Europe. All reported cases were imported from people returning from affected countries. The first laboratory confirmed case was notified in November 2013 in Germany (Tappe et al., 2014). Since then, from 2015, imported cases have been documented in Norway (Waehre et al., 2014), Italy (Zammarchi et al., 2015), and Germany (Tappe et al., 2015). However, after the recent explosive expansion of the virus in the Americas, many European countries have reported imported cases of ZIKV infection in travelers returning home (**Figure 5**), including Austria, Denmark, Finland, France, Germany, Ireland, Italy, Portugal, the Netherlands, Spain, Slovenia, Sweden, Switzerland, and UK[18].

As commented above, currently only *A. albopictus* has colonized Europe, mainly the Mediterranean area (Medlock et al., 2012; Kraemer et al., 2015). Since *A. aegypti* is the main responsible of the current transmission of ZIKV in other regions of the world, the actual risk of a ZIKV outbreak in Europe seems to be low. Nonetheless, beside the possible reintroduction of *A. aegypti* in the continent, *A. albopticus* is circulating in the southern regions of the continent and, since this species

[12]http://www.health.gov.au/internet/main/publishing.nsf/Content/ohp-zika-health-practitioners.htm

[13]http://www.arphs.govt.nz/health-information/communicable-disease/dengue-fever-zika-chikungunya#.VsLrOyuG-WM

[14]http://smartraveller.gov.au/bulletins/zika_virus

[15]http://portalsaude.saude.gov.br/images/pdf/2015/dezembro/09/Microcefalia---Protocolo-de-vigil--ncia-eresposta---vers--o-1----09dez2015-8h.pdf

[16]http://www.cdc.gov/zika/geo/united-states.html

[17]http://www.healthycanadians.gc.ca/publications/diseases-conditions-maladies-affections/risks-zika-virus-risques/index-eng.php

[18]http://ecdc.europa.eu/en/healthtopics/zika_virus_infection/zika-outbreak/Pages/epidemiological-situation.aspx

can also efficiently transmit the virus, surveillance programs should be implemented, mainly during warm seasons. In fact, autochthonous DENV and CHIKV infections in France (La Ruche et al., 2010; Delisle et al., 2015), Croatia (Gjenero-Margan et al., 2011), and Italy (Rezza et al., 2007) have already been documented.

CLINICAL MANIFESTATIONS AND PATHOGENY

Humans

Zika virus infection has been reported to be symptomatic only in around 18% of the cases (Duffy et al., 2009), in which it causes a mild, self-limiting disease with an incubation period of up to 10 days, often mistaken with other arboviral infections like dengue or chikungunya. Clinical manifestations in symptomatic cases resemble that of an influenza-like syndrome (Macnamara, 1954), being the most common symptoms fever, rash, arthralgia, and conjunctivitis; with headache, vomiting, edema, and jaundice being reported less frequently (Zammarchi et al., 2015). Digestive complications (abdominal pain, diarrhea, and constipation), mucous membrane ulcerations (aphthae), and pruritus are rarely observed. The symptoms usually resolve spontaneously after 3–7 days, but arthralgia may persist for up to 1 month (Foy et al., 2011). A post-infection asthenia seems to be also frequent. In any case, severe disease requiring hospitalization has been uncommon until now.

First description of clinical manifestations were reported in 1954 by Macnamara (Macnamara, 1954). In 1956, Bearcroft described the symptoms in an experimentally infected human volunteer (Bearcroft, 1956). A slight generalized headache and a rise in temperature started at the 3rd day, increasing during day 5. By day 7 the patient felt well and the temperature had fallen to normal. Jaundice did not develop, and no other evidence of hepatic dysfunction was found. The total white blood cells during the first 12 days after inoculation did not differ greatly from pre-inoculation counts. Urine estimations were consistently negative for bile pigments and albumin during this period. Estimations of total serum bilirubin carried out between the time of inoculation and the 17th day showed normal levels, and the site of inoculation appeared healthy.

In 1964, Simpson described his own acquired ZIKV illness while working with ZIKV strains isolated from *A. africanus* collected during 1962–1963 (Simpson, 1964). The illness began with a slight frontal headache, showing no other symptoms at the time. During the 2nd day he presented a diffuse pink maculopapular rash which covered the face, neck, trunk, and upper arms that spread gradually to involve all four limbs, and felt slight aching sensations in his back and thighs. Oral temperature at this time was normal in the morning, however he was slightly febrile throughout the day. The temperature returned to normal by the end of the day, and he felt much better, apart from a slight headache. The rash persisted to day 5, when it was fading until completely disappear. No other signs or symptoms were noted during the illness.

Information on laboratory tested alterations associated with ZIKV infection are limited, but may include transient leucopenia, with (Kutsuna et al., 2014) or without thrombocytopenia (Kwong et al., 2013), and slight elevation of serum lactate dehydrogenase, gamma-glutamyl transferase, and of inflammatory parameters (C– reactive protein, fibrinogen and ferritin) (Tappe et al., 2014). Serum aspartate aminotransferase (AST) and alanine aminotransferase (ALT) concentrations may or may not be elevated.

Guillain–Barre Syndrome (GBS)

An association of ZIKV infection with more severe disease outcomes, such as GBS has been also proposed. GBS is an autoimmune disease causing acute or subacute flaccid paralysis that can even cause death (van den Berg et al., 2013; Dominguez-Moreno et al., 2014), and that it has been previously associated with other flaviviral infections (Puccioni-Sohler et al., 2012). Remarkably, as aforementioned, during ZIKV outbreak reported in French Polynesia the incidence rate of GBS cases was approximately 20-fold higher than expected given the size of the French Polynesia population and its previously established incidence (1–2/100 000 population per year) (Oehler et al., 2014). Likewise, in Colombia, during the ongoing outbreak, 86 cases of GBS have been associated to ZIKV infection[19]. Based on the 600 000 expected Zika infections in the country, up to 1000 cases of GBS could be anticipated. These data point to a worrisome increase in the potential clinical severity of the disease (Roth et al., 2014).

Microcephaly

Similarly to GBS, even more disturbing is the astonishing rise in the number of babies born with microcephaly and neurological disorders that have been suggested to be associated with the current ZIKV outbreak in Brazil (Schuler-Faccini et al., 2016). These congenital infections presumably due to ZIKV exposure have been also associated with an increase in vision-threatening findings, which include bilateral macular and perimacular lesions, as well as optic nerve abnormalities in most cases (de Paula Freitas et al., 2016; Ventura et al., 2016a,b). By the end of 2015, Brazil's Health Ministry reported an unusual spike in reported cases of microcephaly in the northeastern state of Pernambuco, where the affected children's mothers had been in early pregnancy at around the same time as large ZIKV outbreaks occurred. The Ministry subsequently raised the alarm of a possible link to ZIKV. Although most cases have been described in Pernambuco, they have also been diagnosed in other eight Brazilian states so far (Oliveira Melo et al., 2016). However, it should be noted that neighboring Colombia has reported over 5000 cases of ZIKV in pregnant women and so far, only one documented case of microcephaly in a newborn, and other two with other congenital brain abnormalities have been very recently documented (Butler, 2016).

[19]http://www.who.int/csr/don/12-february-2016-gbs-colombia-venezuela/en/

In any case, this possible association of microcephaly with ZIKV is still a matter of controversy among researchers. The Latin American Collaborative Study of Congenital Malformations (ECLAMC) suggested that the rise in reported cases of microcephaly might largely be attributable to the intense search for cases of the birth defect, and to misdiagnoses, that arose from heightened awareness in the wake of the possible link with ZIKV. However, it should be noted that a successful RT-PCR amplification of the complete ZIKV genome from a fetal brain tissue has been recently described (Mlakar et al., 2016). In this line, an expectant mother presented a febrile illness with rash at the end of the first trimester of pregnancy while she was living in Brazil. Ultrasonography performed at 29 weeks of gestation revealed microcephaly with calcifications in the fetal brain and placenta. After the mother requested termination of the pregnancy, the fetal autopsy revealed microcephaly, with almost complete agyria, hydrocephalus, and multifocal dystrophic calcifications in the cortex and subcortical white matter, with associated cortical displacement and mild focal inflammation. Electron microscopy analysis also revealed spherical virus particles with morphologic characteristics consistent with ZIKV. Likewise, very recently, the ZIKV genome has been detected and sequenced form amniotic fluid samples of two pregnant women whose fetuses where diagnosed with microcephaly, (Calvet et al., 2016).

The apparent risk of microcephaly was enough for the World Health Organization (WHO) to declare a public health emergency of international concern on February 1 (Rubin et al., 2016) that has lately been integrated into risk assessments by the ECDC[4]. Therefore, the most important current milestone in ZIKV investigations is to clearly elucidate the possible relationship of the infection with the development of serious neurological disorders.

Animal Models

The pathogenicity of ZIKV has been also evaluated in several animal models. Among these, the most widely used has been the mouse model. Moreover, initial isolation of ZIKV was performed by intracerebral inoculation of viral samples into mice (Dick, 1952). On these initial experiments, infectious virus was only recovered from infected brain mice, whereas no infectious virus able to replicate when inoculated in other mice was recovered from non-nervous tissues such as kidney, lung, liver, or spleens, highlighting the marked neurotropism of ZIKV. Infected mice displayed detectable signs of infection about 5–6 days post-infection, time at which virus titers peaked (Dick, 1952). The analysis of the pathological changes observed in ZIKV-infected mice brains revealed various stages of cellular infiltration and degeneration that were also found in the spinal cords. Degeneration of nerve cells, especially in the region of the hippocampus, resulted in an early and marked enlargement of astroglial cells with patchy destruction of the pyriform cells of Ammon's horn (Weinbren and Williams, 1958; Bell et al., 1971). Microscopy analysis confirmed that the virus replicates in both neurons and astrogial cells (Weinbren and Williams, 1958; Bell et al., 1971). Remarkably, while mice of all ages tested were susceptible to intracerebral inoculations, mice of 2 week of age and over could rarely be infected by the intraperitoneal route. In contrast, mice younger than 2 weeks were highly susceptible to intraperitoneal inoculation of the virus (Dick, 1952). This finding is similar to that reported for other flaviviruses, such as Usutu virus (Blazquez et al., 2015). ZIKV-induced pathological changes are confined to the central nervous system in older animals, but myocarditis (and associated pulmonary oedema) and skeletal myositis can be also found in young (1–5 days old) animals infected with ZIKV (Weinbren and Williams, 1958). However, it has to be considered that only non-mouse-adapted strain of ZIKV seems to induce myocarditis (Weinbren and Williams, 1958). Besides mice, the susceptibility to ZIKV has been also evaluated in other non-primate small animal models, demonstrating that cotton-rats, guinea pigs, and rabbits did not show clinical signs of infection after intracerebral inoculation of a late passage mouse brain virus. Nevertheless, inoculation of low passage ZIKV in guinea pigs resulted in death at 6 days post-infection (Dick, 1952).

The pathogenicity of ZIKV in monkeys seems to be also mild, except from the sentinel Rhesus 766 from which ZIKV was initially isolated that exhibited a slight pyrexia (Dick et al., 1952). Experimentally infected monkeys developed an unapparent infection after subcutaneous inoculation with mouse recovered ZIKV. Only one of five monkeys tested showed a mild pyrexia after intracerebral inoculation, whereas the others showed no signs of infection (Dick, 1952). In any case, it should be noted that all infected monkeys showed viremia during the 1st week after inoculation, as well as induction of specific antibodies about 14 days post-inoculation (Dick, 1952). This induction of viremia in monkeys is supposed to play an essential role for the establishment of the enzootic transmission cycle between non-human primates and mosquitoes.

DIAGNOSIS

Different arboviral infections can have similar clinical presentations and, thus, their circulation may be underreported if specific diagnostic tools have not been implemented. In the case of ZIKV, diagnosis presents several drawbacks; there is no "gold standard" diagnosis tool, antibodies are frequently cross-reactive between flaviviruses, which limits the use of serology, viral culture is not routinely performed, and, so far, there is no antigenic detection test available (Musso et al., 2015a). At present, the diagnosis of ZIKV infection is mainly made through "in house" molecular (RT-PCR) and serologic [Ig M ELISA and plaque reduction neutralization test (PRNT)] assays, as only very recently molecular commercial tests for ZIKV have been made available[20,21] .

Antibody

An IgM ELISA was developed at the Arboviral Diagnostic and Reference Laboratory of the CDC to detect ZIKV using samples from Yap Island outbreak (Lanciotti et al., 2008). IgM

[20]http://www.euroimmun.co.uk/recent-news/first-commercial-antibody-tests-for-zika-virus-diagnostics

[21]www.genesig.com/assets/files/zikv.pdf

was detectable as early as 3 days after onset of illness, but cross-reaction with other flaviviruses was observed. This cross-reactivity of sera from convalescent-phase patients was more frequent in those from patients with evidence of previous flavivirus infections than among those with apparent primary ZIKV infections, mainly in the case of previous DENV infection (Lanciotti et al., 2008; Duffy et al., 2009). This IgM cross-reactivity has been further corroborated (Zammarchi et al., 2015; Shinohara et al., 2016), so, as in many other flaviviral infections, ELISA positive results should be confirmed, either by testing an acute-phase serum sample collected as early as possible after onset of illness and a second sample collected 2–3 weeks later, or by additional assays, such as PRNT, that have to result in at least a fourfold increase in ZIKV neutralizing antibody titers when compared with that of the other viruses tested[22].

Recently, a report analyzing the presence of IgG against different flaviviruses among blood donors in French Polynesia, based on the use of recombinant antigens comprising the domain III of the envelope protein of each virus strain, claimed to differentiate between them, as an overall seropositivity rates of 80.3% for at least one DENV serotype, 0.8% for ZIKV, 1.3% for JEV, and 1.5% for WNV were recorded (Aubry et al., 2015). In any case, it should be noticed that ELISA cross-reactivity can also be the result of co-infections with more than one flavivirus that co-circulate in the same regions, as recently reported for two cases (Dupont-Rouzeyrol et al., 2015).

Nucleic Acid

The diagnosis of ZIKV is at present primarily based on the detection of viral RNA from specimens by means of RT-PCR (Faye et al., 2008; Cao-Lormeau et al., 2014; Grard et al., 2014; Musso et al., 2015a; Zammarchi et al., 2015; Marcondes and Ximenes, 2016; Shinohara et al., 2016). However, as the viremic period is short, direct virus detection should be performed in samples took during the first 3–5 days after the onset of symptoms (Balm et al., 2012).

A specific ZIKV nucleic acid test (NAT) was implemented in routine practice during the French Polynesian outbreak (Musso et al., 2014a) on the basis of protocols implemented to prevent WNV transmission by transfusion in North America. Additional NATs have been developed based on specific Asian and African ZIKV strains targeting either the envelope or the NS5 regions (Duffy et al., 2009; Faye et al., 2013). Likewise, a one-step RT-PCR assay to detect ZIKV in human serum has also been developed (Faye et al., 2008). The assay showed to be rapid, sensitive, and specific to detect ZIKV in cell culture or serum. Nevertheless, since experimentally infected samples were used, it needs to be validated for diagnosis using clinical samples.

All these NATs have several limitations. Because of its nucleotide sequence specificity, NAT cannot be used to screen a wide range of pathogens with one run, being necessary the use of multiple assays if several pathogens are co-circulating in the same area, but multiple testing is expensive and time-consuming (Aubry et al., 2016), and it does not detect all infected blood donations, especially when nucleic acid loads are low and when sera are tested in large pools (Musso et al., 2015a). Therefore, use of alternative source of samples has been proposed. In this sense, the suitability of urine samples for diagnosing ZIKV infection has been recently confirmed, as RNA of the virus is detectable in urine at a higher load and with a longer duration than in serum (10 to >20 days, and >7 days once it become undetectable in serum) (Gourinat et al., 2015; Shinohara et al., 2016). Saliva has also been used as an alternative sample for routine ZIKV RNA detection, showing positivity more frequently than blood samples, but it did not increase the window of detection in contrast to what was reported for urine. Since ZIKV RNA detection was found negative in some saliva samples while positive in blood, saliva cannot replace blood samples, but just help on detection (Musso et al., 2015a). As a result of these data, it has been suggested to collect both blood and saliva samples to increase the sensitivity of molecular detection of viral RNA for Zika fever diagnosis in acute phase. On the other hand, urine sample can be associated at the late stages of the disease. Thus, when detection of ZIKV is of particular importance, using a combination of samples (blood/saliva/urine) is recommended (Musso et al., 2015a).

PROPHYLAXIS

Currently there are no specific antiviral agents, vaccine, or prophylaxis for ZIKV. Treatment is generally directed to symptom relief with analgesics and anti-pyretics. Therefore, developing an effective, safe, and affordable ZIKV vaccine, and search for antiviral effective compounds for disease treatment is a current challenge for Zika disease.

Vaccines for various flaviviruses have been produced during the past years, some of them being already in the market, such as those for YFV or WNV (Dauphin and Zientara, 2007; Ulbert and Magnusson, 2014; Monath and Vasconcelos, 2015). These vaccines have been produced using different strategies, inactivated or live-attenuated viruses, recombinant proteins or peptides expressed in different heterologous systems, recombinant subviral particles, chimeric backbone viruses, or naked cDNA (Ishikawa et al., 2014). Thus, it seems reasonable to think that similar strategies can be applied to ZIKV. For instances, a DNA based vaccine (SynCon, Phamaceutical, U.S.) has been recently produced, which it is expected to go into Phase I trials before the end of 2016. In addition, a global patent of two vaccine candidates (a recombinant vaccine and an inactivated vaccine) for ZIKV has been just filled (Bharat Biotech, India).

Developing effective specific therapies for ZIKV seems much more difficult, as so far, and despite several attempts have been made in the past years, none of such compounds are available against any flaviviral infection (Apte-Sengupta et al., 2014). Along with drugs, antibody mediated protection against ZIKV should also be addressed, as in other flaviviral infections (YF, WN, or dengue) experimental specific antibodies treatment has been some time successful (Roehrig et al., 2001; Ben-Nathan et al., 2003; Engle and Diamond, 2003).

Nevertheless, and despite the great effort that probably will be made within the scientific community in the coming years, it

[22]http://www.cdc.gov/mmwr/volumes/65/wr/mm6503e3.htm

will take time until any drugs or vaccines against ZIKV will be commercially available.

Public Health and Preventive Measures

As previously detailed, ZIKV infection generally causes a non-severe disease (Zammarchi et al., 2015), but some areas newly affected by the virus are providing worrisome information on the already mentioned potential association of neonatal malformations (microcephaly) and GBS with ZIKV (Oehler et al., 2014; Mlakar et al., 2016; Tetro, 2016). As a consequence, and in the absence of another explanation for these connections, the WHO declared a Public Health Emergency of International Concern on the 1st of February of 2016[23], highlighting the importance of enhance the measures to reduce the ZIKV infection, particularly among pregnant women and women of childbearing age. In this way, as previously remarked, additional investigations are absolutely necessary to determine whether there is a casual link between ZIKV and microcephaly and GBS.

Currently, no vaccine or treatment exists to prevent ZIKV disease, although more probably several vaccine prototypes will be obtained soon. Thus, to facilitate the surveillance and control measures, research and development efforts should be intensified for ZIKV vaccines, therapeutics, and improved new diagnostics[23], especially if the connection between ZIKV and microcephaly is confirmed.

In the meantime, prevention, and control of ZIKV infection are mainly focused on avoiding the bites of carrying mosquitoes responsible for disease transmission. These measures are the same as those recommended to prevent other diseases transmitted by mosquitoes bite (Martin-Acebes and Saiz, 2012), and include the use of insect repellent, wearing long-sleeved shirts and long pants, the elimination of standing water where mosquitoes can lay eggs, the minimization of outdoor activities coincident with the maximum activity of mosquitoes, the installation of window and door screens, and the implementation of accurate mosquito control programs[24]. ECDC and CDC recommend those pregnant women and those who are trying to get pregnant consider not travel to areas affected by the virus[24,25]. Furthermore, some American countries, including Colombia[26], Honduras[27], Ecuador[28], El Salvador[29], and Jamaica[30] have recommended delay pregnancy during the ZIKV outbreak.

On the other hand, in territories where autochthonous ZIKV infection has not yet been detected, but potential vectors are present, the public health institutions should have a quick response to prevent local transmission when an imported case is confirmed. In this context, WHO called for health authorities to collaborate with the transport sector to ensure disinfection of aircraft from affected areas[23].

As described before, a few cases of sexual transmission of ZIKV has been reported (Musso et al., 2015b; McCarthy, 2016). Even though further studies are needed to confirm the actual sexual risk of ZIKV transmission, the CDC recommend the proper use of preventive measures, such as condoms, if having sex (vaginal, anal, or oral) with a male partner while traveling, or if he has just comeback from an area where the virus is actively circulating; even more, in this latter case, sexual abstinence is recommended if the partner is pregnant[31].

Likewise, and since spread of the virus through blood transfusion has been reported (Musso et al., 2014a), the countries affected by the virus should take measures to prevent this way of infection (Marano et al., 2016). Moreover, in territories free of ZIKV, people returning from affected areas should delay blood donations[32].

The rapid spread of ZIKV in Brazil and the Americas, with the suspected neonatal malformations associated, is being cause for alarm and the mass media are even questioning if the 2016 Olympic Games in Rio should be delayed or even canceled. In fact, the Brazilian authorities are performing intense vector control programs to eliminate mosquito vector that transmit ZIKV and other arboviruses (Petersen et al., 2016).

FINAL REMARKS

After the discovery of ZIKV in 1947, several investigations were performed during the 1950s and 1960s to characterize the new pathogen but, since then, little more was done until the spread of the virus to the Micronesia in 2007. After that, with the explosive invasion of the Americas in 2015, and the possible association of ZIKV with severe neurological diseases, these investigations have significantly increased, and it should be expected that much more information regarding the biology and the clinical consequences of ZIKV infection will be available soon. Several issues have to be resolved. For instance, most of the virus molecular biology and the virus-host cell interactions have been inferred from other related flaviviruses, and, thus, they should be specifically analyzed. Likewise, the elucidation of the presence of determinants of virulence/pathogenicity, and the role that protein glycosylations can play on it will be of much interest. On the other hand, the role of humans in the spread of ZIKV, and whether vertical and sexual transmission routes are important factors for human-to-human spread also need to be clearly determined. No least is the improvement of more accurate diagnostic tools, the development of vaccines, and the design of antiviral therapies. Finally, without any doubt, currently the most important and pressing issue is to reveal whether ZIKV

[23]http://www.who.int/mediacentre/news/statements/2016/1st-emergency-committee-zika/en/

[24]http://www.cdc.gov/zika/prevention/index.html

[25]http://ecdc.europa.eu/en/healthtopics/zika_virus_infection/factsheet-health-professionals/Pages/factsheet_health_professionals.aspx

[26]https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/Circular-02-de-2016.pdf

[27]http://www.salud.gob.hn/

[28]http://www.salud.gob.ec/ministerio-de-salud-refuerza-recomendaciones-a-mujeres-embarazadas-por-virus-zika/

[29]http://www.salud.gob.sv/novedades/noticias/noticias-ciudadanosas/348-enero-2016/3275--25-01-2016-sistema-nacional-de-proteccion-civil-realiza-jornada-municipal-contra-zancudo-transmisor-del-dengue-chikv-y-zika.html

[30]http://moh.gov.jm/presentation/notes-for-minister-of-health-hon-horace-dalley-post-cabinet-press-briefing-january-20-2016-at-11a-m-office-of-the-prime-minister/

[31]http://wwwnc.cdc.gov/travel/notices/alert/zika-virus-central-america

[32]http://www.msssi.gob.es/gabinete/notasPrensa.do?id=3900

infection is the cause of the increase in the number of GBS and microcephaly cases lately reported.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by grant ZIKA-BIO from INIA to J-CS, and AGL2014-56518-JIN from MINECO to MM-A, AV-C is a recipient of a "Contrato de formación postodoctoral" from MINECO. TM-R is a recipient of a "Formación de Personal Investigador (FPI)" pre-doctoral fellowship from INIA.

REFERENCES

Acosta, E. G., Kumar, A., and Bartenschlager, R. (2014). Revisiting dengue virus-host cell interaction: new insights into molecular and cellular virology. *Adv. Virus Res.* 88, 1–109. doi: 10.1016/B978-0-12-800098-4.00001-5

Adams, S. C., Broom, A. K., Sammels, L. M., Hartnett, A. C., Howard, M. J., Coelen, R. J., et al. (1995). Glycosylation and antigenic variation among Kunjin virus isolates. *Virology* 206, 49–56. doi: 10.1016/S0042-6822(95)80018-2

Akoua-Koffi, C., Diarrassouba, S., Benie, V. B., Ngbichi, J. M., Bozoua, T., Bosson, A., et al. (2001). [Investigation surrounding a fatal case of yellow fever in Cote d'Ivoire in 1999]. *Bull. Soc. Pathol. Exot.* 94, 227–230.

Alera, M. T., Hermann, L., Tac-An, I. A., Klungthong, C., Rutvisuttinunt, W., Manasatienkij, W., et al. (2015). Zika virus infection, Philippines, 2012. *Emerg. Infect. Dis.* 21, 722–724. doi: 10.3201/eid2104.141707

Althouse, B. M., Hanley, K. A., Diallo, M., Sall, A. A., Ba, Y., Faye, O., et al. (2015). Impact of climate and mosquito vector abundance on sylvatic arbovirus circulation dynamics in Senegal. *Am. J. Trop. Med. Hyg.* 92, 88–97. doi: 10.4269/ajtmh.13-0617

Apte-Sengupta, S., Sirohi, D., and Kuhn, R. J. (2014). Coupling of replication and assembly in flaviviruses. *Curr. Opin. Virol.* 9, 134–142. doi: 10.1016/j.coviro.2014.09.020S1879-6257(14)00202-8

Aubry, M., Richard, V., Green, J., Broult, J., and Musso, D. (2016). Inactivation of Zika virus in plasma with amotosalen and ultraviolet A illumination. *Transfusion* 56, 33–40. doi: 10.1111/trf.13271

Aubry, M., Teissier, A., Roche, C., Richard, V., Yan, A. S., Zisou, K., et al. (2015). Chikungunya outbreak, French Polynesia, 2014. *Emerg. Infect. Dis.* 21, 724–726. doi: 10.3201/eid2104.141741

Balm, M. N., Lee, C. K., Lee, H. K., Chiu, L., Koay, E. S., and Tang, J. W. (2012). A diagnostic polymerase chain reaction assay for Zika virus. *J. Med. Virol.* 84, 1501–1505. doi: 10.1002/jmv.23241

Baronti, C., Piorkowski, G., Charrel, R. N., Boubis, L., Leparc-Goffart, I., and de Lamballerie, X. (2014). Complete coding sequence of zika virus from a French polynesia outbreak in 2013. *Genome Announc* 2, e500-14. doi: 10.1128/genomeA.00500-14

Bearcroft, W. G. (1956). Zika virus infection experimentally induced in a human volunteer. *Trans. R. Soc. Trop. Med. Hyg.* 50, 442–448. doi: 10.1016/0035-9203(56)90090-6

Bell, T. M., Field, E. J., and Narang, H. K. (1971). Zika virus infection of the central nervous system of mice. *Arch. Gesamte Virusforsch* 35, 183–193. doi: 10.1007/BF01249709

Ben-Nathan, D., Lustig, S., Tam, G., Robinzon, S., Segal, S., and Rager-Zisman, B. (2003). Prophylactic and therapeutic efficacy of human intravenous immunoglobulin in treating West Nile virus infection in mice. *J. Infect. Dis.* 188, 5–12. doi: 10.1086/376870

Berthet, F. X., Zeller, H. G., Drouet, M. T., Rauzier, J., Digoutte, J. P., and Deubel, V. (1997). Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. *J. Gen. Virol.* 78(Pt 9), 2293–2297. doi: 10.1099/0022-1317-78-9-2293

Berthet, N., Nakoune, E., Kamgang, B., Selekon, B., Descorps-Declere, S., Gessain, A., et al. (2014). Molecular characterization of three Zika flaviviruses obtained from sylvatic mosquitoes in the Central African Republic. *Vector Borne Zoonotic Dis.* 14, 862–865. doi: 10.1089/vbz.2014.1607

Besnard, M., Lastere, S., Teissier, A., Cao-Lormeau, V., and Musso, D. (2014). Evidence of perinatal transmission of Zika virus, French Polynesia, December 2013 and February 2014. *Euro. Surveill.* 19:20751. doi: 10.2807/1560-7917.ES2014.19.13.20751

Blazquez, A. B., Escribano-Romero, E., Martin-Acebes, M. A., Petrovic, T., and Saiz, J. C. (2015). Limited susceptibility of mice to Usutu virus (USUV) infection and induction of flavivirus cross-protective immunity. *Virology* 482, 67–71. doi: 10.1016/j.virol.2015.03.020

Blazquez, A. B., Escribano-Romero, E., Merino-Ramos, T., Saiz, J. C., and Martin-Acebes, M. A. (2014). Stress responses in flavivirus-infected cells: activation of unfolded protein response and autophagy. *Front. Microbiol.* 5:266. doi: 10.3389/fmicb.2014.00266

Buathong, R., Hermann, L., Thaisomboonsuk, B., Rutvisuttinunt, W., Klungthong, C., Chinnawirotpisan, P., et al. (2015). Detection of Zika Virus Infection in Thailand, 2012-2014. *Am. J. Trop. Med. Hyg.* 93, 380–383. doi: 10.4269/ajtmh.15-0022

Buckley, A., and Gould, E. A. (1988). Detection of virus-specific antigen in the nuclei or nucleoli of cells infected with Zika or Langat virus. *J. Gen. Virol.* 69(Pt 8), 1913–1920. doi: 10.1099/0022-1317-69-8-1913

Butler, D. (2016). First Zika-linked birth defects detected in Colombia. *Nature* 531:153. doi: 10.1038/nature.2016.19502

Calvet, G. A., Filippis, A. M., Mendonca, M. C., Sequeira, P. C., Siqueira, A. M., Veloso, V. G., et al. (2016). First detection of autochthonous Zika virus transmission in a HIV-infected patient in Rio de Janeiro. *Brazil. J. Clin. Virol.* 74, 1–3. doi: 10.1016/j.jcv.2015.11.014

Campos, G. S., Bandeira, A. C., and Sardi, S. I. (2015). Zika Virus outbreak, Bahia, Brazil. *Emerg. Infect. Dis.* 21, 1885–1886. doi: 10.3201/eid2110.150847

Cao-Lormeau, V. M., Roche, C., Teissier, A., Robin, E., Berry, A. L., Mallet, H. P., et al. (2014). Zika virus, French polynesia, South pacific, 2013. *Emerg. Infect. Dis.* 20, 1085–1086. doi: 10.3201/eid2006.140138

Caron, M., Paupy, C., Grard, G., Becquart, P., Mombo, I., Nso, B. B., et al. (2012). Recent introduction and rapid dissemination of Chikungunya virus and Dengue virus serotype 2 associated with human and mosquito coinfections in Gabon, central Africa. *Clin. Infect. Dis.* 55, e45–e53. doi: 10.1093/cid/cis530cis530

Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., Errett, J., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452–456. doi: 10.1038/nature09489

Darwish, M. A., Hoogstraal, H., Roberts, T. J., Ahmed, I. P., and Omar, F. (1983). A sero-epidemiological survey for certain arboviruses (Togaviridae) in Pakistan. *Trans. R. Soc. Trop. Med. Hyg.* 77, 442–445. doi: 10.1016/0035-9203(83)90108-6

Dauphin, G., and Zientara, S. (2007). West Nile virus: recent trends in diagnosis and vaccine development. *Vaccine* 25, 5563–5576. doi: 10.1016/j.vaccine.2006.12.005

de Paula Freitas, B., de Oliveira Dias, J. R., Prazeres, J., Sacramento, G. A., Ko, A. I., Maia, M., et al. (2016). Ocular findings in infants with microcephaly associated with presumed zika virus congenital infection in salvador, brazil. *JAMA Ophthalmol.* doi: 10.1001/jamaophthalmol.2016.0267 [Epub ahead of print].

Delatte, H., Desvars, A., Bouetard, A., Bord, S., Gimonneau, G., Vourc'h, G., et al. (2010). Blood-feeding behavior of *Aedes albopictus*, a vector of chikungunya on la reunion. *Vector Borne Zoonotic Dis.* 10, 249–258. doi: 10.1089/vbz.2009.0026

Delisle, E., Rousseau, C., Broche, B., Leparc-Goffart, I., L'Ambert, G., Cochet, A., et al. (2015). Chikungunya outbreak in Montpellier, France, September to October 2014. *Euro Surveill.* 20:21108. doi: 10.2807/1560-7917.ES2015.20.17.21108

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi: 10.1093/nar/gkn180

Diagne, C. T., Diallo, D., Faye, O., Ba, Y., Gaye, A., Dia, I., et al. (2015). Potential of selected senegalese *Aedes* spp. mosquitoes (Diptera: Culicidae) to transmit Zika virus. *BMC Infect. Dis.* 15:492. doi: 10.1186/s12879-015-1231-2

Diallo, D., Sall, A. A., Diagne, C. T., Faye, O., Ba, Y., Hanley, K. A., et al. (2014). Zika virus emergence in mosquitoes in southeastern Senegal, 2011. *PLoS ONE* 9:e109442. doi: 10.1371/journal.pone.0109442

Dick, G. W. (1952). Zika virus. II. Pathogenicity and physical properties. *Trans. R. Soc. Trop. Med. Hyg.* 46, 521–534. doi: 10.1016/0035-9203(52)90043-6

Dick, G. W., Kitchen, S. F., and Haddow, A. J. (1952). Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* 46, 509–520. doi: 10.1016/0035-9203(52)90042-4

Dominguez-Moreno, R., Tolosa-Tort, P., Patino-Tamez, A., Quintero-Bauman, A., Collado-Frias, D. K., Miranda-Rodriguez, M. G., et al. (2014). [Mortality associated with a diagnosis of Guillain-Barre syndrome in adults of Mexican health institutions]. *Rev. Neurol.* 58, 4–10.

Dong, H., Fink, K., Zust, R., Lim, S. P., Qin, C. F., and Shi, P. Y. (2014). Flavivirus RNA methylation. *J. Gen. Virol.* 95, 763–778. doi: 10.1099/vir.0.062208-0

Duffy, M. R., Chen, T. H., Hancock, W. T., Powers, A. M., Kool, J. L., Lanciotti, R. S., et al. (2009). Zika virus outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* 360, 2536–2543. doi: 10.1056/NEJMoa0805715

Dupont-Rouzeyrol, M., O'Connor, O., Calvez, E., Daures, M., John, M., Grangeon, J. P., et al. (2015). Co-infection with Zika and dengue viruses in 2 patients, New Caledonia, 2014. *Emerg. Infect. Dis.* 21, 381–382. doi: 10.3201/eid2102.141553

Dyer, O. (2016). Sixty seconds on, Zika virus. *BMJ* 352:467. doi: 10.1136/bmj.i467

Engle, M. J., and Diamond, M. S. (2003). Antibody prophylaxis and therapy against West Nile virus infection in wild-type and immunodeficient mice. *J. Virol.* 77, 12941–12949. doi: 10.1128/JVI.77.24.12941-12949.2003

Fagbami, A. H. (1979). Zika virus infections in Nigeria: virological and seroepidemiological investigations in Oyo State. *J. Hyg.* 83, 213–219. doi: 10.1017/S0022172400025997

Fauci, A. S., and Morens, D. M. (2016). Zika virus in the americas–yet another arbovirus threat. *N. Engl. J. Med.* 374, 601–604. doi: 10.1056/NEJMp1600297

Faye, O., Diallo, D., Diallo, M., Weidmann, M., and Sall, A. A. (2013). Quantitative real-time PCR detection of Zika virus and evaluation with field-caught mosquitoes. *Virol. J.* 10:311. doi: 10.1186/1743-422X-10-311

Faye, O., Dupressoir, A., Weidmann, M., Ndiaye, M., and Alpha Sall, A. (2008). One-step RT-PCR for detection of Zika virus. *J. Clin. Virol.* 43, 96–101. doi: 10.1016/j.jcv.2008.05.005

Faye, O., Freire, C. C., Iamarino, A., de Oliveira, J. V., Diallo, M., Zanotto, P. M., et al. (2014). Molecular evolution of Zika virus during its emergence in the 20(th) century. *PLoS Negl. Trop. Dis.* 8:e2636. doi: 10.1371/journal.pntd.0002636

Filipe, A. R., Martins, C. M., and Rocha, H. (1973). Laboratory infection with Zika virus after vaccination against yellow fever. *Arch. Gesamte Virusforsch* 43, 315–319. doi: 10.1007/BF01556147

Foy, B. D., Kobylinski, K. C., Chilson Foy, J. L., Blitvich, B. J., Travassos da Rosa, A., Haddow, A. D., et al. (2011). Probable non-vector-borne transmission of Zika virus, Colorado, USA. *Emerg. Infect. Dis.* 17, 880–882. doi: 10.3201/eid1705.101939

Freire, C. C. D. M., Iamarino, A., Neto, D. F. D. L., Sall, A. A., and Zanotto, P. M. D. A. (2015). Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. *bioRxiv* doi: 10.1101/032839

Gatherer, D., and Kohl, A. (2016). Zika virus: a previously slow pandemic spreads rapidly through the Americas. *J. Gen. Virol.* 97, 269–273. doi: 10.1099/jgv.0.000381

Geser, A., Henderson, B. E., and Christensen, S. (1970). A multipurpose serological survey in Kenya. 2. Results of arbovirus serological tests. *Bull. World Health Organ.* 43, 539–552.

Gjenero-Margan, I., Aleraj, B., Krajcar, D., Lesnikar, V., Klobucar, A., Pem-Novosel, I., et al. (2011). Autochthonous dengue fever in Croatia, August-September 2010. *Euro Surveill.* 16:19805.

Gollins, S. W., and Porterfield, J. S. (1986). The uncoating and infectivity of the flavivirus West Nile on interaction with cells: effects of pH and ammonium chloride. *J. Gen. Virol.* 67(Pt 9), 1941–1950. doi: 10.1099/0022-1317-67-9-1941

Gould, E. A., and Solomon, T. (2008). Pathogenic flaviviruses. *Lancet* 371, 500–509. doi: 10.1016/S0140-6736(08)60238-X

Gourinat, A. C., O'Connor, O., Calvez, E., Goarant, C., and Dupont-Rouzeyrol, M. (2015). Detection of Zika virus in urine. *Emerg. Infect. Dis.* 21, 84–86. doi: 10.3201/eid2101.140894

Grard, G., Caron, M., Mombo, I. M., Nkoghe, D., Mboui Ondo, S., Jiolle, D., et al. (2014). Zika virus in Gabon (Central Africa)–2007: a new threat from *Aedes albopictus*? *PLoS Negl. Trop. Dis.* 8:e2681. doi: 10.1371/journal.pntd.0002681

Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., et al. (2012). Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl. Trop. Dis.* 6:e1477. doi: 10.1371/journal.pntd.0001477

Haddow, A. J., and Dick, G. W. (1948). Catches of biting Diptera in Uganda, with anaesthetized monkeys as bait. *Ann. Trop. Med. Parasitol.* 42, 271–277. doi: 10.1080/00034983.1948.11685376

Haddow, A. J., Williams, M. C., Woodall, J. P., Simpson, D. I., and Goma, L. K. (1964). Twelve isolations of Zika virus from *Aedes* (*Stegomyia*) Africanus (Theobald) Taken in and above a Uganda Forest. *Bull. World Health Organ.* 31, 57–69.

Hamel, R., Dejarnac, O., Wichit, S., Ekchariyawat, P., Neyret, A., Luplertlop, N., et al. (2015). Biology of Zika virus infection in human skin cells. *J. Virol.* 89, 8880–8896. doi: 10.1128/JVI.00354-15

Hammon, W. M., Schrack, W. D. Jr., and Sather, G. E. (1958). Serological survey for a arthropod-borne virus infections in the Philippines. *Am. J. Trop. Med. Hyg.* 7, 323–328.

Heang, V., Yasuda, C. Y., Sovann, L., Haddow, A. D., Travassos da Rosa, A. P., Tesh, R. B., et al. (2012). Zika virus infection, Cambodia, 2010. *Emerg. Infect. Dis.* 18, 349–351. doi: 10.3201/eid1802.111224

Higgs, S. (2016). Zika virus: emergence and emergency. *Vector Borne Zoonotic Dis.* 16, 75–76. doi: 10.1089/vbz.2016.29001.hig

Holmes, E. C. (2003). Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. *J. Virol.* 77, 11296–11298. doi: 10.1128/JVI.77.20.11296-11298.2003

Ishikawa, T., Yamanaka, A., and Konishi, E. (2014). A review of successful flavivirus vaccines and the problems with those flaviviruses for which vaccines are not yet available. *Vaccine* 32, 1326–1337. doi: 10.1016/j.vaccine.2014.01.040S0264-410X(14)00074-7

Jan, C., Languillat, G., Renaudet, J., and Robin, Y. (1978). [A serological survey of arboviruses in Gabon]. *Bull. Soc. Pathol. Exot. Filiales* 71, 140–146.

Johnson, B. K., Chanas, A. C., Shockley, P., Squires, E. J., Gardner, P., Wallace, C., et al. (1977). Arbovirus isolations from, and serological studies on, wild and domestic vertebrates from Kano Plain, Kenya. *Trans. R. Soc. Trop. Med. Hyg.* 71, 512–517. doi: 10.1016/0035-9203(77)90146-8

Kraemer, M. U., Sinka, M. E., Duda, K. A., Mylne, A. Q., Shearer, F. M., Barker, C. M., et al. (2015). The global distribution of the arbovirus vectors *Aedes aegypti* and *A. albopictus*. *eLife* 4:e08347. doi: 10.7554/eLife.08347

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., et al. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell* 108, 717–725. doi: 10.1016/S0092-8674(02)00660-8

Kuno, G., and Chang, G. J. (2007). Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. *Arch. Virol.* 152, 687–696. doi: 10.1007/s00705-006-0903-z

Kuno, G., Chang, G. J., Tsuchiya, K. R., Karabatsos, N., and Cropp, C. B. (1998). Phylogeny of the genus Flavivirus. *J. Virol.* 72, 73–83.

Kutsuna, S., Kato, Y., Takasaki, T., Moi, M., Kotaki, A., Uemura, H., et al. (2014). Two cases of Zika fever imported from French Polynesia to Japan, December 2013 to January 2014 [corrected]. *Euro Surveill.* 19:20683. doi: 10.2807/1560-7917.ES2014.19.4.20683

Kwong, J. C., Druce, J. D., and Leder, K. (2013). Zika virus infection acquired during brief travel to Indonesia. *Am. J. Trop. Med. Hyg.* 89, 516–517. doi: 10.4269/ajtmh.13-0029

La Ruche, G., Souares, Y., Armengaud, A., Peloux-Petiot, F., Delaunay, P., Despres, P., et al. (2010). First two autochthonous dengue virus infections in metropolitan France, September 2010. *Euro Surveill.* 15:19676.

Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Velez, J. O., Lambert, A. J., Johnson, A. J., et al. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* 14, 1232–1239. doi: 10.3201/eid1408.080287

Lee, V. H., and Moore, D. L. (1972). Vectors of the 1969 yellow fever epidemic on the Jos Plateau, Nigeria. *Bull. World Health Organ.* 46, 669–673.

Leung, G. H., Baird, R. W., Druce, J., and Anstey, N. M. (2015). Zika virus infection in australia following a monkey bite in indonesia. *Southeast Asian J. Trop. Med. Public Health* 46, 460–464.

Li, J., Bhuvanakantham, R., Howe, J., and Ng, M. L. (2006). The glycosylation site in the envelope protein of West Nile virus (Sarafend) plays an important role in replication and maturation processes. *J. Gen. Virol.* 87, 613–622. doi: 10.1099/vir.0.81320-0

Li, M. I., Wong, P. S., Ng, L. C., and Tan, C. H. (2012). Oral susceptibility of Singapore *Aedes* (*Stegomyia*) aegypti (*Linnaeus*) to Zika virus. *PLoS Negl. Trop. Dis.* 6:e1792. doi: 10.1371/journal.pntd.0001792PNTD-D-12-00044

Lok, S. M. (2016). The interplay of dengue virus morphological diversity and human antibodies. *Trends Microbiol.* 24, 284–293. doi: 10.1016/j.tim.2015.12.004

Luo, D., Vasudevan, S. G., and Lescar, J. (2015). The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res.* 118, 148–158. doi: 10.1016/j.antiviral.2015.03.014

Macnamara, F. N. (1954). Zika virus: a report on three cases of human infection during an epidemic of jaundice in Nigeria. *Trans. R. Soc. Trop. Med. Hyg.* 48, 139–145. doi: 10.1016/0035-9203(54)90006-1

Marano, G., Pupella, S., Vaglio, S., Liumbruno, G. M., and Grazzini, G. (2016). Zika virus and the never-ending story of emerging pathogens and transfusion medicine. *Blood Transfus.* 14, 95–100. doi: 10.2450/2015.0066-15

Marchette, N. J., Garcia, R., and Rudnick, A. (1969). Isolation of Zika virus from *Aedes aegypti* mosquitoes in Malaysia. *Am. J. Trop. Med. Hyg.* 18, 411–415.

Marcondes, C. B., and Ximenes, M. F. (2016). Zika virus in Brazil and the danger of infestation by *Aedes* (*Stegomyia*) mosquitoes. *Rev. Soc. Bras. Med. Trop.* doi: 10.1590/0037-8682-0220-2015 [Epub ahead of print].

Martin-Acebes, M. A., Blazquez, A. B., Jimenez de Oya, N., Escribano-Romero, E., and Saiz, J. C. (2011). West nile virus replication requires Fatty Acid synthesis but is independent on phosphatidylinositol-4-phosphate lipids. *PLoS ONE* 6:e24970. doi: 10.1371/journal.pone.0024970

Martin-Acebes, M. A., and Saiz, J. C. (2012). West Nile virus: a re-emerging pathogen revisited. *World J. Virol.* 1, 51–70. doi: 10.5501/wjv.v1.i2.51

Martines, R. B., Bhatnagar, J., Keating, M. K., Silva-Flannery, L., Muehlenbachs, A., Gary, J., et al. (2016). Notes from the field: evidence of zika virus infection in brain and placental tissues from two congenitally infected newborns and two fetal losses - Brazil, 2015. *Morb. Mortal. Wkly. Rep.* 65, 159–160. doi: 10.15585/mmwr.mm6506e1

McCarthy, M. (2016). Zika virus was transmitted by sexual contact in Texas, health officials report. *BMJ* 352, i720. doi: 10.1136/bmj.i720

McCrae, A. W., and Kirya, B. G. (1982). Yellow fever and Zika virus epizootics and enzootics in Uganda. *Trans. R. Soc. Trop. Med. Hyg.* 76, 552–562. doi: 10.1016/0035-9203(82)90161-4

Medlock, J. M., Hansford, K. M., Schaffner, F., Versteirt, V., Hendrickx, G., Zeller, H., et al. (2012). A review of the invasive mosquitoes in Europe: ecology, public health risks, and control options. *Vector Borne Zoonotic Dis.* 12, 435–447. doi: 10.1089/vbz.2011.0814

Miorin, L., Romero-Brey, I., Maiuri, P., Hoppe, S., Krijnse-Locker, J., Bartenschlager, R., et al. (2013). Three-dimensional architecture of tick-borne encephalitis virus replication sites and trafficking of the replicated RNA. *J. Virol.* 87, 6469–6481. doi: 10.1128/JVI.03456-12

Mlakar, J., Korva, M., Tul, N., Popovic, M., Poljsak-Prijatelj, M., Mraz, J., et al. (2016). Zika Virus associated with microcephaly. *N. Eng. J. Med.* 374, 951–958. doi: 10.1056/NEJMoa1600651

Monath, T. P., and Vasconcelos, P. F. (2015). Yellow fever. *J. Clin. Virol.* 64, 160–173. doi: 10.1016/j.jcv.2014.08.030S1386-6532(14)00369-2

Monath, T. P., Wilson, D. C., Lee, V. H., Stroh, G., Kuteyi, K., and Smith, E. A. (1973). The 1970 yellow fever epidemic in Okwoga District, Benue Plateau State, Nigeria. I. Epidemiological observations. *Bull. World Health Organ.* 49, 113–121.

Monlun, E., Zeller, H., Le Guenno, B., Traore-Lamizana, M., Hervy, J. P., Adam, F., et al. (1993). [Surveillance of the circulation of arbovirus of medical interest in the region of eastern Senegal]. *Bull. Soc. Pathol. Exot.* 86, 21–28.

Mukhopadhyay, S., Kim, B. S., Chipman, P. R., Rossmann, M. G., and Kuhn, R. J. (2003). Structure of West Nile virus. *Science* 302, 248. doi: 10.1126/science.1089316

Mukhopadhyay, S., Kuhn, R. J., and Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. *Nat. Rev. Microbiol.* 3, 13–22. doi: 10.1038/nrmicro1067

Musso, D. (2015). Zika virus transmission from french polynesia to Brazil. *Emerg. Infect. Dis.* 21:1887. doi: 10.3201/eid2110.151125

Musso, D., Nhan, T., Robin, E., Roche, C., Bierlaire, D., Zisou, K., et al. (2014a). Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. *Euro Surveill.* 19:20761. doi: 10.2807/1560-7917.ES2014.19.14.20761

Musso, D., Nilles, E. J., and Cao-Lormeau, V. M. (2014b). Rapid spread of emerging Zika virus in the Pacific area. *Clin. Microbiol. Infect.* 20, O595–O596. doi: 10.1111/1469-0691.12707

Musso, D., Roche, C., Nhan, T. X., Robin, E., Teissier, A., and Cao-Lormeau, V. M. (2015a). Detection of Zika virus in saliva. *J. Clin. Virol.* 68, 53–55. doi: 10.1016/j.jcv.2015.04.021S1386-6532(15)00133-X

Musso, D., Roche, C., Robin, E., Nhan, T., Teissier, A., and Cao-Lormeau, V. M. (2015b). Potential sexual transmission of Zika virus. *Emerg. Infect. Dis.* 21, 359–361. doi: 10.3201/eid2102.141363

Noble, C. G., and Shi, P. Y. (2012). Structural biology of dengue virus enzymes: towards rational design of therapeutics. *Antiviral Res.* 96, 115–126. doi: 10.1016/j.antiviral.2012.09.007

Oehler, E., Watrin, L., Larre, P., Leparc-Goffart, I., Lastere, S., Valour, F., et al. (2014). Zika virus infection complicated by Guillain-Barre syndrome–case report, French Polynesia, December 2013. *Euro Surveill.* 19:20720. doi: 10.2807/1560-7917.ES2014.19.9.20720

Oliveira Melo, A. S., Malinger, G., Ximenes, R., Szejnfeld, P. O., Alves Sampaio, S., and Bispo de Filippis, A. M. (2016). Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: tip of the iceberg? *Ultrasound Obstet Gynecol.* 47, 6–7. doi: 10.1002/uog.15831

Olson, J. G., Ksiazek, T. G., Gubler, D. J., Lubis, S. I., Simanjuntak, G., Lee, V. H., et al. (1983). A survey for arboviral antibodies in sera of humans and animals in Lombok, Republic of Indonesia. *Ann. Trop. Med. Parasitol.* 77, 131–137.

Olson, J. G., Ksiazek, T. G., Suhandiman, and Triwibowo. (1981). Zika virus, a cause of fever in Central Java, Indonesia. *Trans. R. Soc. Trop. Med. Hyg.* 75, 389–393. doi: 10.1016/0035-9203(81)90100-0

Paupy, C., Delatte, H., Bagny, L., Corbel, V., and Fontenille, D. (2009). *Aedes albopictus*, an arbovirus vector: from the darkness to the light. *Microbes Infect.* 11, 1177–1185. doi: 10.1016/j.micinf.2009.05.005S1286-4579(09)00105-1

Petersen, E., Wilson, M. E., Touch, S., McCloskey, B., Mwaba, P., Bates, M., et al. (2016). Rapid spread of Zika virus in the Americas – implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int. J. Infect. Dis.* 44, 11–15. doi: 10.1016/j.ijid.2016.02.001

Pierson, T. C., and Diamond, M. S. (2012). Degrees of maturity: the complex structure and biology of flaviviruses. *Curr. Opin. Virol.* 2, 168–175. doi: 10.1016/j.coviro.2012.02.011

Plevka, P., Battisti, A. J., Junjhon, J., Winkler, D. C., Holdaway, H. A., Keelapang, P., et al. (2011). Maturation of flaviviruses starts from one or more icosahedrally independent nucleation centres. *EMBO Rep.* 12, 602–606. doi: 10.1038/embor.2011.75

Pond, W. L. (1963). Arthropod-borne virus antibodies in sera from residents of South-East Asia. *Trans. R. Soc. Trop. Med. Hyg.* 57, 364–371. doi: 10.1016/0035-9203(63)90100-7

Ponlawat, A., and Harrington, L. C. (2005). Blood feeding patterns of *Aedes aegypti* and *Aedes albopictus* in Thailand. *J. Med. Entomol.* 42, 844–849. doi: 10.1603/0022-2585(2005)042[0821:ISOAAA]2.0.CO;2

Puccioni-Sohler, M., Orsini, M., and Soares, C. N. (2012). Dengue: a new challenge for neurology. *Neurol. Int.* 4:e15. doi: 10.4081/ni.2012.e15ni.2012.e15

Pupella, S., Pisani, G., Cristiano, K., Catalano, L., and Grazzini, G. (2013). West Nile virus in the transfusion setting with a special focus on Italian preventive measures adopted in 2008-2012 and their impact on blood safety. *Blood Transfus.* 11, 563–574. doi: 10.2450/2013.0077-132013.0077-13

Reiter, P. (2010). Yellow fever and dengue: a threat to Europe? *Euro Surveill.* 15:19509.

Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A. C., Panning, M., et al. (2007). Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet* 370, 1840–1846. doi: 10.1016/S0140-6736(07)61779-6

Robin, Y., and Mouchet, J. (1975). [Serological and entomological study on yellow fever in Sierra Leone]. *Bull. Soc. Pathol. Exot. Fil.* 68, 249–258.

Roby, J. A., Setoh, Y. X., Hall, R. A., and Khromykh, A. A. (2015). Post-translational regulation and modifications of flavivirus structural proteins. *J. Gen. Virol.* 96, 1551–1569. doi: 10.1099/vir.0.000097vir.0.000097

Roehrig, J. T., Staudinger, L. A., Hunt, A. R., Mathews, J. H., and Blair, C. D. (2001). Antibody prophylaxis and therapy for flavivirus encephalitis infections. *Ann. N. Y. Acad. Sci.* 951, 286–297. doi: 10.1111/j.1749-6632.2001.tb02704.x

Roth, A., Mercier, A., Lepers, C., Hoy, D., Duituturaga, S., Benyon, E., et al. (2014). Concurrent outbreaks of dengue, chikungunya and Zika virus infections - an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012-2014. *Euro Surveill.* 19:20929. doi: 10.2807/1560-7917.ES2014.19.41.20929

Rubin, E. J., Greene, M. F., and Baden, L. R. (2016). Zika virus and microcephaly. *N. Engl. J. Med.* doi: 10.1056/NEJMe1601862 [Epub ahead of print].

Saluzzo, J. F., Gonzalez, J. P., Herve, J. P., and Georges, A. J. (1981). [Serological survey for the prevalence of certain arboviruses in the human population of the south-east area of Central African Republic (author's transl)]. *Bull. Soc. Pathol. Exot. Fil.* 74, 490–499.

Schuler-Faccini, L., Ribeiro, E. M., Feitosa, I. M., Horovitz, D. D., Cavalcanti, D. P., Pessoa, A., et al. (2016). Possible association between zika virus infection and microcephaly - Brazil, 2015. *Morb. Mortal. Wkly. Rep.* 65, 59–62. doi: 10.15585/mmwr.mm6503e2

Scott, T. W., and Takken, W. (2012). Feeding strategies of anthropophilic mosquitoes result in increased risk of pathogen transmission. *Trends Parasitol.* 28, 114–121. doi: 10.1016/j.pt.2012.01.001S1471-4922(12)00004-9

Shinohara, K., Kutsuna, S., Takasaki, T., Moi, M. L., Ikeda, M., Kotaki, A., et al. (2016). Zika fever imported from Thailand to Japan, and diagnosed by PCR in the urines. *J. Travel Med.* 23:tav011. doi: 10.1093/jtm/tav011tav011

Shirato, K., Miyoshi, H., Goto, A., Ako, Y., Ueki, T., Kariwa, H., et al. (2004). Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus. *J. Gen. Virol.* 85, 3637–3645. doi: 10.1099/vir.0.80247-0

Simpson, D. I. (1964). Zika virus Infection in Man. *Trans. R. Soc. Trop. Med. Hyg.* 58, 335–338. doi: 10.1016/0035-9203(64)90201-9

Smithburn, K. C. (1952). Neutralizing antibodies against certain recently isolated viruses in the sera of human beings residing in East Africa. *J. Immunol.* 69, 223–234.

Smithburn, K. C. (1954). Neutralizing antibodies against arthropod-borne viruses in the sera of long-time residents of Malaya and Borneo. *Am. J. Hyg.* 59, 157–163.

Smithburn, K. C., Kerr, J. A., and Gatne, P. B. (1954). Neutralizing antibodies against certain viruses in the sera of residents of India. *J. Immunol.* 72, 248–257.

Staples, J. E., Dziuban, E. J., Fischer, M., Cragan, J. D., Rasmussen, S. A., Cannon, M. J., et al. (2016). Interim guidelines for the evaluation and testing of infants with possible congenital zika virus infection - United States, 2016. *Morb. Mortal. Wkly. Rep.* 65, 63–67. doi: 10.15585/mmwr.mm6503e3

Stiasny, K., Fritz, R., Pangerl, K., and Heinz, F. X. (2011). Molecular mechanisms of flavivirus membrane fusion. *Amino Acids* 41, 1159–1163. doi: 10.1007/s00726-009-0370-4

Tappe, D., Nachtigall, S., Kapaun, A., Schnitzler, P., Gunther, S., and Schmidt-Chanasit, J. (2015). Acute Zika virus infection after travel to Malaysian Borneo, September 2014. *Emerg. Infect. Dis.* 21, 911–913. doi: 10.3201/eid2105.141960

Tappe, D., Rissland, J., Gabriel, M., Emmerich, P., Gunther, S., Held, G., et al. (2014). First case of laboratory-confirmed Zika virus infection imported into Europe, November 2013. *Euro Surveill.* 19:20685. doi: 10.2807/1560-7917.ES2014.19.4.20685

Tetro, J. A. (2016). Zika and microcephaly: causation, correlation, or coincidence? *Microb. Infect.* 18, 167–168. doi: 10.1016/j.micinf.2015.12.010

Thomas, S. M., Obermayr, U., Fischer, D., Kreyling, J., and Beierkuhnlein, C. (2012). Low-temperature threshold for egg survival of a post-diapause and non-diapause European aedine strain, *Aedes albopictus* (Diptera: Culicidae). *Parasit. Vectors* 5:100. doi: 10.1186/1756-3305-5-100

Tognarelli, J., Ulloa, S., Villagra, E., Lagos, J., Aguayo, C., Fasce, R., et al. (2016). A report on the outbreak of Zika virus on Easter Island, South Pacific, 2014. *Arch. Virol.* 161, 665–668. doi: 10.1007/s00705-015-2695-5

Ulbert, S., and Magnusson, S. E. (2014). Technologies for the development of West Nile virus vaccines. *Future Microbiol.* 9, 1221–1232. doi: 10.2217/fmb.14.67

van den Berg, B., Bunschoten, C., van Doorn, P. A., and Jacobs, B. C. (2013). Mortality in guillain-barre syndrome. *Neurology* 80, 1650–1654. doi: 10.1212/WNL.0b013e3182904fcc

Vazquez-Calvo, A., Saiz, J. C., McCullough, K. C., Sobrino, F., and Martin-Acebes, M. A. (2012). Acid-dependent viral entry. *Virus Res.* 167, 125–137. doi: 10.1016/j.virusres.2012.05.024

Ventura, C. V., Maia, M., Bravo-Filho, V., Gois, A. L., and Belfort, R. Jr. (2016a). Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet* 387:228. doi: 10.1016/S0140-6736(16)00006-4

Ventura, C. V., Maia, M., Ventura, B. V., Linden, V. V., Araujo, E. B., Ramos, R. C., et al. (2016b). Ophthalmological findings in infants with microcephaly and presumable intra-uterus Zika virus infection. *Arq. Bras. Oftalmol.* 79, 1–3. doi: 10.5935/0004-2749.20160002

Vorndam, V., Mathews, J. H., Barrett, A. D., Roehrig, J. T., and Trent, D. W. (1993). Molecular and biological characterization of a non-glycosylated isolate of St Louis encephalitis virus. *J. Gen. Virol.* 74(Pt 12), 2653–2660. doi: 10.1099/0022-1317-74-12-2653

Waehre, T., Maagard, A., Tappe, D., Cadar, D., and Schmidt-Chanasit, J. (2014). Zika virus infection after travel to Tahiti, December 2013. *Emerg. Infect. Dis.* 20, 1412–1414. doi: 10.3201/eid2008.140302

Way, J. H., Bowen, E. T., and Platt, G. S. (1976). Comparative studies of some African arboviruses in cell culture and in mice. *J. Gen. Virol.* 30, 123–130. doi: 10.1099/0022-1317-30-1-123

Weinbren, M. P., and Williams, M. C. (1958). Zika virus: further isolations in the Zika area, and some studies on the strains isolated. *Trans. R. Soc. Trop. Med. Hyg.* 52, 263–268. doi: 10.1016/0035-9203(58)90085-3

Weissenbock, H., Hubalek, Z., Bakonyi, T., and Nowotny, N. (2010). Zoonotic mosquito-borne flaviviruses: worldwide presence of agents with proven pathogenicity and potential candidates of future emerging diseases. *Vet. Microbiol.* 140, 271–280. doi: 10.1016/j.vetmic.2009.08.025

Welsch, S., Miller, S., Romero-Brey, I., Merz, A., Bleck, C. K., Walther, P., et al. (2009). Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe* 5, 365–375. doi: 10.1016/j.chom.2009.03.007

Winkler, G., Heinz, F. X., and Kunz, C. (1987). Studies on the glycosylation of flavivirus E proteins and the role of carbohydrate in antigenic structure. *Virology* 159, 237–243. doi: 10.1016/0042-6822(87)90460-0

Wolfe, N. D., Kilbourn, A. M., Karesh, W. B., Rahman, H. A., Bosi, E. J., Cropp, B. C., et al. (2001). Sylvatic transmission of arboviruses among Bornean orangutans. *Am. J. Trop. Med. Hyg.* 64, 310–316.

Wong, P. S., Li, M. Z., Chong, C. S., Ng, L. C., and Tan, C. H. (2013). *Aedes* (*Stegomyia*) albopictus (Skuse): a potential vector of Zika virus in Singapore. *PLoS Negl. Trop. Dis.* 7:e2348. doi: 10.1371/journal.pntd.0002348PNTD-D-13-00504

Zammarchi, L., Stella, G., Mantella, A., Bartolozzi, D., Tappe, D., Gunther, S., et al. (2015). Zika virus infections imported to Italy: clinical, immunological and virological findings, and public health implications. *J. Clin. Virol.* 63, 32–35. doi: 10.1016/j.jcv.2014.12.005

Zanluca, C., de Melo, V. C., Mosimann, A. L., Dos Santos, G. I., Dos Santos, C. N., and Luz, K. (2015). First report of autochthonous transmission of Zika virus in Brazil. *Mem. Inst. Oswaldo Cruz* 110, 569–572. doi: 10.1590/0074-02760150192

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Saiz, Vázquez-Calvo, Blázquez, Merino-Ramos, Escribano-Romero and Martín-Acebes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 19 December 2016
doi: 10.3389/fmicb.2016.02028

Development of a Zika Virus Infection Model in Cynomolgus Macaques

*Fusataka Koide[1], Scott Goebel[1], Beth Snyder[1], Kevin B. Walters[1], Alison Gast[1], Kimberly Hagelin[1], Raj Kalkeri[1] and Jonathan Rayner[2]**

[1] Department of Infectious Disease Research, Drug Development, Southern Research Institute, Frederick, MD, USA,
[2] Department of Infectious Disease Research, Drug Development, Southern Research, Birmingham, AL, USA

Limited availability of Indian rhesus macaques (IRM) is a bottleneck to study Zika virus (ZIKV) pathogenesis and evaluation of appropriate control measures in non-human primates. To address these issues, we report here the Mauritian cynomolgus macaque (MCM) model for ZIKV infection. In brief, six MCMs (seronegative for Dengue and ZIKV) were subdivided into three cohorts with a male and female each and challenged with different doses of Asian [PRVABC59 (Puerto Rico) or FSS13025 (Cambodia)] or African (IBH30656) lineage ZIKV isolates. Clinical signs were monitored; and biological fluids (serum, saliva, and urine) and tissues (testes and brain) were assessed for viral load by quantitative reverse transcription polymerase chain reaction and neutralizing antibodies (Nab) by 50% Plaque Reduction Neutralization Test ($PRNT_{50}$) at various times post-infection (p.i). PRVABC59 induced viremia detectable up to day 10, with peak viral load at 2–3 days p.i. An intermittent viremia spike was observed on day 30 with titers reaching 2.5×10^3 genomes/mL. Moderate viral load was observed in testes, urine and saliva. In contrast, FSS13025 induced viremia lasting only up to 6 days and detectable viral loads in testes but not in urine and saliva. Recurrent viremia was detected but at lower titers compare to PRVABC59. Challenge with either PRVABC59 or FSS13025 resulted in 100% seroconversion; with mean $PRNT_{50}$ titers ranging from 597 to 5179. IBH30656 failed to establish infection in MCM suggesting that MCM are susceptible to infection with ZIKV isolates of the Asian lineage but not from Africa. Due to the similarity of biphasic viremia and Nab responses between MCM and IRM models, MCM could be a suitable alternative for evaluation of ZIKV vaccine and therapeutic candidates.

Keywords: Zika virus, cynomolgus macaque, non-human primate, flavivirus, arbovirus

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Akatsuki Saito,
Osaka University, Japan
Anu Susan Charles,
Louisiana State University, USA

****Correspondence:***
Jonathan Rayner
jrayner@southresearch.org

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *21 October 2016*
Accepted: *02 December 2016*
Published: *19 December 2016*

Citation:
Koide F, Goebel S, Snyder B, Walters KB, Gast A, Hagelin K, Kalkeri R and Rayner J (2016) Development of a Zika Virus Infection Model in Cynomolgus Macaques. Front. Microbiol. 7:2028. doi: 10.3389/fmicb.2016.02028

INTRODUCTION

Zika virus (ZIKV) is a flavivirus transmitted primarily through mosquitos, first reported in 1947 (reviewed in Plourde and Bloch, 2016). Though it was initially restricted to Africa and Asia, in the last couple of years increasing cases have been observed in as many as 70 countries worldwide. Though apparent clinical symptoms are reported only in 20% of the infected patients, the disease associated complications such as microcephaly in newborn children and neurological manifestations (Guillain–Barre syndrome) make ZIKV a major health concern. Considering the rapid spread and the associated disease complications, the World Health Organization (WHO) has recently declared ZIKV as a global health emergency. This health scare is compounded by the lack of effective prophylactic and therapeutic measures.

Currently, immune compromised mouse (Brault et al., 2016; Cugola et al., 2016; Dowall et al., 2016; Lazear et al., 2016; Zmurko et al., 2016) and rhesus macaque models (Abbink et al., 2016; Dudley et al., 2016) have been used for studies on the natural history and pathogenesis of ZIKV infection. Type-I interferon receptor deficient AG129 mice but not the parent 129Sv/Ev strain of mice were found to be susceptible to a lethal ZIKV infection (Dowall et al., 2016). Lazear et al. (2016) reported the development of neurological disease in IFNar1 (−/−) mice and IRF3, 5, and 7 triple knockout mice (Lazear et al., 2016). The AG129 model was also helpful in evaluating the antiviral activity of viral polymerase inhibitor 7-deaza-2′-C-methyladenosine (7DMA) (Zmurko et al., 2016). Using the Swiss Jim Lambert (SJL) mouse model, Cugola et al. (2016) were able to demonstrate fetal infection and microcephaly with a Brazilian strain of Zika virus. Though mouse models are easily accessible, non-human primates (NHPs) are an attractive model for ZIKV research and drug discovery due to their close similarity with humans. NHP models could provide invaluable information regarding mechanism of action, efficacy and safety of both drug and vaccine candidates and allow optimization of the product, dose and route as observed previously for HIV vaccines (Spearman, 2006). Rhesus macaques were shown to be susceptible to an Asian lineage of ZIKV (Dudley et al., 2016), with pregnant animals being viremic for longer period compared to non-pregnant animals. Evaluation of three vaccine candidates in rhesus monkeys successfully protected them against ZIKV challenge (Abbink et al., 2016).

The use of a number of non-rhesus macaque species, especially cynomolgus macaques, as a model for human infectious diseases has increased in recent years (Antony and MacDonald, 2015). This is mostly due to the reduced availability of Rhesus monkeys consequent to the ban on their export from India. Compared to Rhesus macaques, Cynomolgus macaques offer the advantages of smaller size and weight (Andrade et al., 2004), resulting in reduced amounts of drugs needed for studies administered on body weight basis. Smaller animal size also provides the additional benefit of easier animal husbandry practices (such as handling, space requirements, etc.), translating into significant cost-benefit. Considering these factors, we conducted a limited study ($N = 2$/group) to evaluate the suitability of cynomolgus monkeys as a potential alternative NHP model for ZIKV infection. Using a systematic approach of infection with ZIKV strains of different geographical origin, we demonstrate that cynomolgus monkeys can be successfully infected with ZIKV of Asian-lineage including isolates recently emerging in the current pandemic of the Americas, but not strains of African lineage.

MATERIALS AND METHODS

Care and Use of Animals

This study was designed to use the fewest number of animals possible, consistent with the objective of the study, the scientific needs, contemporary scientific standards, and in consideration of applicable regulatory requirements. This study design was reviewed by the IACUC at Southern Research Institute and was approved on 04/21/2016; it was assigned IACUC tracking number 16-03-014F. Animals were socially housed during the quarantine phase and single housed following the Day 0 challenge. Animals were housed in stainless steel cages that meet requirements as set forth in the Animal Welfare Act (Public Law 99-198) and the *Guide for the Care and Use of Laboratory Animals* (8th Edition, Institute of Animal Resources, Commission on Life Sciences, National Research Council; National Academy Press; Washington D.C.; 2011). Animals were housed in an environmentally monitored and ventilated room. Fluorescent lighting provided illumination approximately 12 h per day. The objective of this pilot proof of concept study was to evaluate the susceptibility of cynomolgus macaques to ZIKV of different geographic origins and did not involve statistical comparison between groups of animals. We've selected two monkeys per challenge strain and this is deemed sufficient to provide enough data to monitor immunological and virological endpoints against each strain. The use of two animals per strain is the minimum number sufficient to achieve the research goals. Based on the results obtained from this pilot study, statistically relevant sample size will be determined for future GLP studies.

Orchiectomy Surgery

The animals were initially given either atropine (0.02–0.04 mg/kg IM) to control respiratory secretions, then sedated with ketamine (10–50 mg/kg IM). Ketamine was followed by xylazine (0.30 mg/kg IM) for induction. The animals were then intubated, placed on a portable isoflurane machine, and isoflurane (0.5–5.0%) was used to bring the animals to the desired plane of anesthesia for the procedure. Anesthesia was maintained using approximately 1–3% isoflurane throughout the procedure. Before the initial incision, ketoprofen (2.2 mg/kg IM, SID) and buprenorphine (0.01–0.03 mg/kg IM, BID) were administered. A lidocaine/bupivacaine local block (at no more than 1.0 mg/kg of each agent) was given at the incision area before surgery began. Orchiectomy was performed to one testes per day. After surgery, animals were removed from the isoflurane machine and placed on towels/blankets with a warming system (Bair-Hugger mat) and monitored until they recovered their swallowing reflex. At this time, the endotracheal tube was removed and the animal was returned to its home cage. Animals were observed continually by trained technicians until they were able to sit upright, at which time they were considered to have recovered from the anesthesia. Ketoprofen (2.2 mg/kg IM, SID) and buprenorphine (0.01–0.03 mg/kg IM, BID) were given as post-operative analgesia for at least 2 days following the day of surgery.

Sequence Analysis

Comparative ZIKV genomic sequence analysis and alignments were performed using the NCBI BLAST Suite[1]. The "Megablastn, Blastn" programs were used for the genomic comparisons and "Primer-blast" for sequence specific primer homology analysis.

[1]http://www.ncbi.nlm.nih.gov/pubmed

Viruses and Cell Culture

Vero cells (ATCC, CCL81) were grown in Dulbecco's Minimal Essential Medium (DMEM, Lonza), supplemented with 10% fetal bovine serum (FBS), NEAA and L-Glutamine according to standard culture conditions. ZIKV strains were obtained from the following sources: IBH30656 (BEI Resources), PRVABC59 (CDC, Division of Vector-borne Infectious Diseases) and FSS13025 (UTMB Arbovirus Reference Collection). In addition, the following ZIKV isolates were used as reference strains for comparative sequence analysis: ZKV2015 (Genbank Accession #KU497555.1) (Calvet et al., 2016), Brazilian isolate Natal RGN, Bahia, Brazil (Seq. Accession# KU527068.1) (Calvet et al., 2016; Mlakar et al., 2016). Viral stocks of each strain were amplified in Vero cells and quantitated according to the standard plaque assay methodology.

Primers and Probes

Polymerase chain reaction (PCR) primers were designed to the NS1 region of the ZIKV genome, proximal to those described (Lanciotti et al., 2008) with modified target sequences enabling the detection of isolates PRVABC59 Puerto Rico (Asian lineage) and FSS13025 Cambodia (Asian lineage), IBH30656 Nigeria (African lineage). Primer and probe sequences are provided in Goebel et al. (2016). Primer and probe sequences were characterized for compatible melting temperatures (Tm), self-dimer and hairpin potential using the Integrated DNA Technologies (IDT, Coralville, IA) "Primer Analysis software" tool. All the primers and probes were synthesized at IDT Technologies. All primers were subjected to sequence analysis including specific ZIKV strain homology and hybridization potential across different clades. The probe has a 5′ 6-FAM reporter and an internal (9th position) ZEN quencher and a 3′ IBFQ Iowa black quencher. Primers and protocols used for the generation of the RNA template used for the standard curve for absolute quantitation were as described in Goebel et al. (2016).

Quantitative RT-PCR

Viral RNA was extracted from biological fluid samples (serum, urine, and saliva) and tissue samples, including testes and brain. Blood samples for serum were collected from anesthetized animals on 1–4, 6, 8, 10, 14, 30, and 60 days post-infection (p.i). Urine and oral swab saliva samples were collected on 1–4, 6, 8, 10, 14 days p.i. Urine samples were collected from anesthetized animals by catheterization or cystocentesis. Urine was not successfully collected at all time points, especially from females (indicated in **Table 1**). Saliva samples were obtained by swabbing the oral cavity with a cotton swab, followed by emersion of the swab into 2 mL PBS without Calcium and Magnesium (Lonza, 17-512F). Briefly, using the QIAmp Viral RNA Mini kit (Qiagen, 52906) total RNA was extracted and purified from a total biological sample volume of 140 μl (as per manufacturer) and eluted into 60 μL of nuclease free water (Ambion, AM9939). Total RNA was extracted from tissues samples as per manufacturer using the RNeasy Lipid Tissue mini kit (Qiagen, 74804). Briefly, several (100–150 mg) sections were dissected from each tissue sample. Tissue sections were subjected to mechanical tissue homogenization in 1 mL of QIAzol Lysis Reagent (Qiagen, 1023537) using a handheld tissue homogenizer (Omni International) fitted with a rotor stator. Purified RNA was eluted in 50 μL nuclease free water.

Five μl of purified RNA from each test article was used in a 20 μl qRT-PCR reaction consisting of Fast Virus 4x Master Mix (Applied Biosystems, 4444436) containing 500 nM forward and reverse primers and 200 nM probe. Cycling parameters included an initial reverse transcription (RT) for 5 min at 53°C, followed by 1 min at 95°C and 45 cycles of two step cycling at 95°C for 5 s and 60°C for 50 s. A standard curve using positive control RNA template was established over a dynamic range of 6-logs (10^6–10^1) with each dilution in triplicate for absolute quantitation of test samples. The Ct values obtained for each of the 6 log dilutions indicate each strain specific primer probe set to be consistently within 1 Ct of each other. Furthermore, each of the strain specific primer and probe sets for the qRT-PCR reaction were found to have a very similar lower limit of quantitation (LLOQ) with values of no more than 10 copies/reaction or 500 copies/mL (data not shown).

Statistics

The mean viral genome copies found in 4 and 8 day p.i. teste samples were compared by one-way analysis of variance (ANOVA) followed by Bonferroni's multiple comparison test. Differences in means were considered statistically significant for a p-value < 0.05.

Plaque Reduction Neutralization Test ($PRNT_{50}$)

Briefly, Vero cells seeded at a concentration of approximately 3×10^5 cells/mL in 24-well plates were incubated for approximately 24 h. Day of assay, input virus (PRVABC59) and serially diluted serum samples were mixed and incubated for 1 h at 37 ± 1°C in the dilution plate. Supernatant from cell-seeded 24-well plates was decanted, then 100 μl of virus/serum mixture was transferred from the dilution plate to the cells. After 1 h adsorption, agarose-containing overlay media was added and plates were incubated at 37 ± 1°C, 5% CO_2 for 3 days. The cells were fixed and stained using crystal violet solution and plaques were counted visually. The neutralizing antibody titer was expressed as the highest test serum dilution for which the virus infectivity is reduced by 50%.

RESULTS

Cohorts

A total of six cynomolgus macaques, seronegative for Dengue and ZIKV by ELSIA were subdivided in three cohorts, each containing a male and female. Animals were inoculated subcutaneously with ZIKV from either the Asian lineage (Puerto Rico isolate; PRVABC59, or Cambodian isolate; FSS13025) or an African lineage (Nigerian isolate; IBH30656). Male and female animals were inoculated on Day 0 with 5.0×10^5 PFU and

TABLE 1 | Virus shedding in urine and saliva.

	Urine						Oral Swabs					
	PRVABC59		FSS13025		IBH30656		PRVABC59		FSS13025		IBH30656	
Day	#5262M	#5258F	#5260M	#5259F	#5261M	#5257F	#5262M	#5258F	#5260M	#5259	#5261M	#5257F
1	0	0	0	N/C	0	N/C	0	0	0	1.6	N/T	N/T
2	265	0	N/C	N/C	43	N/C	105	0	0	0	N/T	N/T
3	0	0	19	N/C	N/C	N/C	198	0	52	0	N/T	N/T
4	0	N/C	N/C	N/C	0	N/C	169	146	0	1	N/T	N/T
6	0	29	0	N/C	0	N/C	0	0	6.6	2	N/T	N/T
8	0	0	N/C	N/C	0	N/C	138	142	0	81	N/T	N/T
10	0	87	0	N/C	200	N/C	244	0	12	0	N/T	N/T
14	0	114	N/C	N/C	0	N/C	109	359	0	0	N/T	N/T

N/C, not collected; N/T, not tested.

1.0×10^4 PFU, respectively. Previous studies have shown viremia in cynomologus macaques following infection with similar concentrations of Dengue virus. Additionally, natural levels of West Nile Virus (WNV) transmission through a mosquito "bite" reportedly range between 1.0×10^4 and 1.0×10^6 PFU (Styer et al., 2007; Cox et al., 2012). Biological fluids (serum, salvia, and urine) were collected at various times between 1 and 60 days p.i. Selected tissue samples were also collected including testes on days 4 and 8 and brain tissue upon animal sacrifice at study termination on day 60. Samples were assessed for viral load and shedding using reverse transcription followed by quantitative PCR (qRT-PCR) and immunological responses were assessed using the 50% Plaque Reduction Neutralization Test ($PRNT_{50}$).

Clinical Observations

One animal challenged with IBH30656 developed a slight erythema around the injection site (Draize dermal score of 1) on study days 8 and 10. The erythema had resolved by day 14. No clinical symptoms were observed in any of the other animals during the course of the study.

Serum Viral Load Analysis

Viral load was determined by qRT-PCR in serum recovered from blood collected on days 1–4, 6, 8, 10, 14, 30, and 60. Both cohorts challenged with Asian lineage strains, PRVABC59 or FSS13025, resulted in substantial viral load that peaked between 2 and 3 days p.i. Maximal serum load for the PRVABC59 challenge (high/low dose) was (1.4×10^5/8.9×10^3) genome copies/mL and for FSS13025 (1.7×10^4/1.8×10^4) copies/mL (**Figure 1**). However, the cohort infected with the African lineage strain IBH30656 failed to produce a "productive" infection as viral load peaked 3 day p.i. at less than 100 copies/mL, well below our defined LLOQ detection.

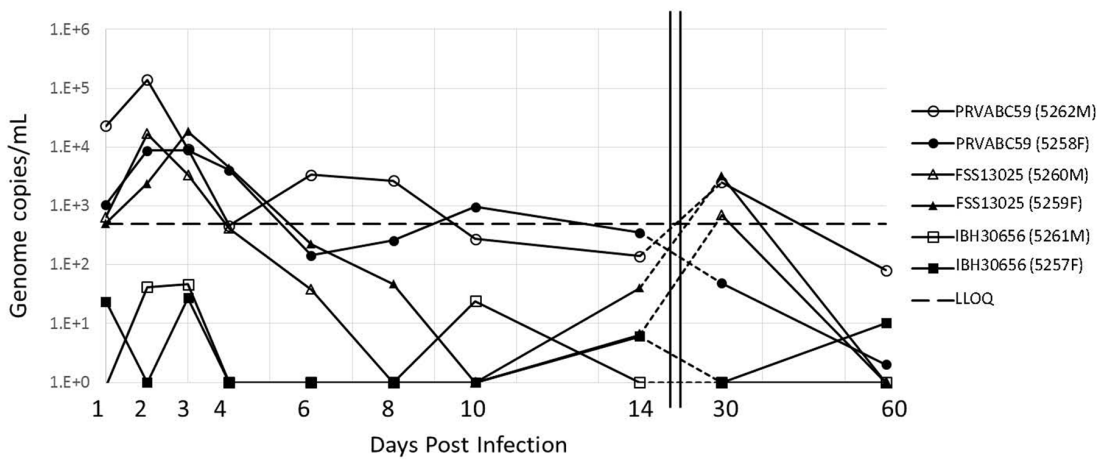


FIGURE 1 | Viral load in the serum of ZIKV infected cynomolgus macaques. Animals were grouped into three cohorts, each containing a male (opened symbols) and female (closed symbols). Each cohort was challenged with ZIKV isolates at the dose indicated with either PRVABC59 (●), FSS13025 (▲), or IBH30656 (■). The serum viral load was determined using qRT-PCR at various days p.i. Notably, the serum viral load peaks around day 2–3 p.i. The LLOQ was previously determined to be 500 copies/mL and is indicated by the black dashed line.

Viral Clearance

Shortly after peaking at 2 or 3 days p.i., serum viremia diminished over the course of about 10 days depending on the virus. By 14 days p.i. all viral loads were below the LLOQ of the qRT-PCR assay at 500 copies/mL. Interestingly, we observed a consistent viral load rebound of about 2 log by 30 day p.i. in three of the four animals infected with either of the two Asian lineage strains. This viral load rebound from serum samples is consistent with that previously reported from a ZIKV challenge in rhesus macaques (Dudley et al., 2016). In cynomolgus macaques the rebound seems to persist longer, at least through day 30 p.i. which was the only intermediate viral load test point between day 14 and the study termination at day 60. Viremia associated with the later viral load rebound diminished to less than 100 copies/mL by 60 days p.i. (**Figure 1**).

Viral Shedding in Body Fluids

In contrast to the robust amounts of virus found in the serum of animals challenged with the Asian lineage viruses (PRVABC59 and FSS13025), saliva and urine samples produced moderate levels of viral load (<300 copies/mL). Lower viral loads in saliva may be partly due to the dilution of swabs in PBS prior to analysis. As expected, considering the serum viral load, the animals challenged with the African strain IBH30656 resulted in extremely low, sporadic amounts of detectable viral RNA levels (**Table 1**).

Tissue Analysis

One testicle was harvested at 4 and 8 days p.i. from the male of each cohort. The highest viral loads were consistently detected from testes samples collected on day 8 post-inoculation. Interestingly, the FSS13025 challenged male had the highest testes viral load (**Figure 2**). Again consistent with the serum analysis, no viral load was detected in the testes of animal challenged with the African strain IBH30656. Upon study termination (60 days p.i.) total brain samples were harvested from all animals. Multiple tissue sections from each brain were tested for viral load by qRT-PCR and all were negative for any residual viral genome copies.

Neutralizing Antibody Response

Neutralizing antibodies (Nabs) following Zika challenge are illustrated in **Table 2**. Sera collected on day 0 prior to challenge were all negative ($PRNT_{50} \leq 10$). The IBH30656 challenged monkeys (5261 and 5257), that exhibited very low viral load throughout study period, also failed to produce cross reactive Nab to PRVABC59. The monkeys inoculated with the PRVABC59 and FSS13025 achieved high Nab titer to PRVABC59 by day 14. Monkey 5262 that recorded highest viremia also attained highest Nab titer of 7920. The day 14 Nab titers for these monkeys ranged from 1,053 to 7,920. All four monkeys suffered a moderate drop in titer on day 30 but Nab activity was persistent and achieved titers ranging from 797 to 2,380. Notably, monkey 5262 (PRVABC59) and 5259 (FSS13025), that experienced more apparent viremia rebound by 30 days p.i., attained >2-fold higher Nab response than the other monkey in the same cohort during this time. Nab titer was sustainable and all four macaques achieved durable $PRNT_{50}$ titers ranging from 528 ≥ 1,071 at 60 days p.i.

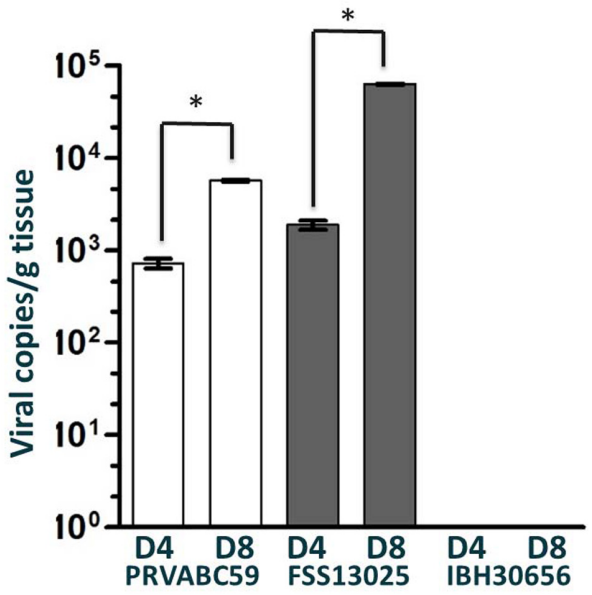


FIGURE 2 | Viral load in the testes of ZIKV infected cynomolgus macaques. Viral load in the testes of cynomolgus macaques infected with PRVABC59, FSS13025, or IBH30656 were detected by qRT-PCR on days 4 and 8 p.i. Multiple sections of each testicle were tested independently and reported here are the sections with the highest viral load. *p-value < 0.05.

DISCUSSION

Arthropod-borne viruses such as the flaviviruses continue to pose a significant threat to global health. The origin and global spread of ZIKV during the ongoing pandemic is well-documented (Dick et al., 1952; Duffy et al., 2009; Haddow et al., 2012; Musso et al., 2014; Campos et al., 2015; Zanluca et al., 2015; Solomon et al., 2016). The significance of ZIKV is highlighted in the infection associated complications, such as Guillain–Barre syndrome (Oehler et al., 2014; do Rosario et al., 2016) and the risk of vertical transmission of the virus from mother to the fetus, which can result in devastating lifelong neurological complications including microcephaly (Brasil and Nielsen-Saines, 2016; Martines et al., 2016; Mlakar et al., 2016; Sarno et al., 2016). Considering the global health emergency posed by ZIKV, there is an ongoing concerted effort among the scientific community to develop new diagnostics, vaccines and antivirals to stem the pandemic. Identification of cost effective and robust animal models amenable to measuring viral load, viral shedding and building cellular and humoral immune correlates of vaccine efficacy is critical for the development of new vaccines and therapeutics. Here, we report the results from a limited study demonstrating the utility of cynomolgus macaques as a cost effective alternative NHP animal model for the study of ZIKV infection of Asian lineage.

TABLE 2 | Development of anti-zika neutralizing antibodies following challenge.

Animal ID	Challenge strain	Dose (PFU)	$PRNT_{50}$ Titer			
			D2	D14	D30	D60
5262 M	PRVABC59	5.0×10^5	<10	7920	2380	1071
5258 F	PRVABC59	1.0×10^4	<10	2438	875	740
5260 M	FSS13025	5.0×10^5	<10	1053	797	528
5259 F	FSS13025	1.0×10^4	<10	3131	2019	666
5261 M	IBH 30656	5.0×10^5	<10	<10	<10	<10
5257 F	IBH 30656	1.0×10^4	<10	<10	<10	<10

Anti-PRVABC59 neutralizing antibodies were measured in serum collected at 2, 14, 30, and 60 days p.i. by standard PRNT assay. PRNT end-point titers are expressed as the reciprocal of the last serum dilution showing the 50% reduction in plaque counts ($PRNT_{50}$).

Failure of the African strain IBH30656 to produce a significant viremia is in contrast to the previous work done in rhesus macaques by the Wisconsin National Primate Research Center[2], where rhesus macaques developed a robust viremia when challenged with the prototypic African strain, Uganda MR766 (GenBank, LC002520, obtained from CDC, Ft. Collins, CO, USA). In addition to the differences in animal models, potentially significant genetic differences between the two African lineage strains may exist. These differences may encode changes leading to altered codon utilization or post-translation processing which have been proposed to play a role in ZIKV pathogenicity (Haddow et al., 2012; Gupta et al., 2016; van Hemert and Berkhout, 2016). Indeed, these two isolates have differences in passage history and share only 93% identity at the nucleotide level. Heterogeneity between available MR766 sequences makes genetic comparisons with IbH30656 difficult (Haddow et al., 2012). Additional studies are needed to further characterize the host and viral factors involved in divergent pathogenicity between ZIKV strains.

Interestingly, the Puerto Rican isolate PRVABC59 appeared to provide the most robust infection, resulting in rapid systemic serum viral load within the first 24 h p.i. (2.3×10^4 and 1.1×10^3 copies/mL) compared to the Cambodian FSS13025 isolate (6.4×10^2 and 5.0×10^2 copies/mL). The temporary lag in amplification of the FSS13025 isolate may suggest different replication kinetics between these two strains. In the rhesus macaque model, peak plasma viremia occurred between 2 and 6 days p.i. and peak virus titer ranged from 8.2×10^4 to 2.5×10^6 RNA copies/mL after 1.0×10^4 to 1.0×10^6 challenge with Asian strains (Dudley et al., 2016). Altogether, our study shows the peak viral load was moderately lower in the cynomolgus macaque, the duration and overall kinetics of virus replication produced by Asian strains were similar to those observed in the rhesus monkeys.

Recent bioinformatics and phylogenetic analyses, have suggested the possible derivation of the currently circulating (in the Americas) ZIKV pandemic isolates from the Asian lineage (Brasil et al., 2016; Calvet et al., 2016; Lanciotti et al., 2016). Furthermore, the sequences of clinical samples isolated from the American outbreak are highly conserved, and are therefore now recognized as a new clade of the Asian lineage (Enfissi et al., 2016; Wang et al., 2016; Ye et al., 2016). Using genomic sequence alignments, a comparison between historic Asian lineage isolates, currently circulating American clade isolates and original African lineage isolates demonstrates how divergent the American clade has become from the African lineage (**Table 3**). In our study, infection of NHPs with the African lineage virus (IBH30656) did not elicit an immune response that cross-neutralized to PRVABC59. This may be due to antigenicity difference between African and Asian lineage viruses but most likely is due to the failure of the virus to establish a productive infection. Numerous studies have now been published on comparative genomics and epidemiological analysis aimed at identifying critical adaptive mutation(s) or processes such as codon utilization or glycosylation that may contribute to the pathogenicity and/or fitness in transmission of the emergent American clade (Gupta et al., 2016; van Hemert and Berkhout, 2016; Ye et al., 2016).

[2]https://zika.labkey.com/project/OConnor/ZIKV-002/begin.view

TABLE 3 | Genomic sequence comparison of ZIKV isolates.

ZIKV Strain: Lineage/clade:	*MR-766 African*	*IBH30656 African*	*FSS13025 Asian*	*PRVABC59 Asian*	*ZKV2015 American*
MR-766	----	93	89	89	89
IBH30656	93	----	88	88	88
FSS13025	89	88	----	98	98
PRVABC59	89	88	98	----	99
ZKV2015	89	88	98	99	----

Pairwise comparison of nucleotide sequence using NCBI (https://blast.ncbi.nlm.nih.gov) Blastn sequence alignment tool. Strains of African and Asian lineage were aligned with the newly emerging American Clade (Ye et al., 2016). ZKV2015 (Accession #KU497555.1) represents the American clade, the most clinically relevant strains circulating in Brazil.

Broad systemic infection of cynomolgus macaques with two independent Asian lineage ZIKV isolates including one from the current circulating pandemic, observed in our studies, suggest the clinical relevance of this model. The failure to detect systemic infection after challenge with the the Nigerian IBH30656 strain of African lineage, is confounding, as other strains of the African lineage (MR766) have been used in rhesus macaques and found to produce robust infections in that animal model as stated above. Whether this observation is limited to Nigerian IBH30656 or more broadly, to other African isolates in cynomolgus macaques need to be explored further.

Additionally, it would be of great interest to elucidate the nature of the late recurrent viremia, now consistently seen in both the cynomolgus and rhesus models reported previously (Dudley et al., 2016). Although this "rebound virus" is eventually cleared in the NHP animal models discussed, the identification of tissue specific reservoir(s) for the rebound virus could lead to mechanisms posing additional adaptive risks and changes in the progression of the infection in humans, either through enhanced transmission or potentially escalating clinical complications.

In the context of viral reservoirs, it is interesting to note that the viral load of both PRVABC59 and FSS13025 in the testes at 8 days p.i. were higher in comparison to 4 days p.i. Importantly, this is well after the serum viral load had peaked between 2 and 3 days p.i. This observation is consistent with previously reported that male anatomical tissues have provided safe harbor for persistent viral titers, long after symptoms and or systemic viral load have resolved (Osuna et al., 2016).

Interestingly, Dudley et al. (2016) suggests that the viral plasma "blips" seen after systemic clearance maybe the result of viral seeding from these reservoirs. Importantly, understanding tissue specific viral reservoirs and their role in providing for mechanisms of viral adaptive immune escape would be critical to the understanding of non-arthropod vectored transmission in the current pandemic strain, and may provide strategic insights for the development of new vaccines and therapeutics.

In this study, detection of serum and tissue viral load was associated with the development of high functional Nab titers in cynomolgus macaques infected with PRVABC59. These results provide insight into the possibility of cynomolgus macaques as an alternative model to study vaccine immunogenicity and efficacy to decipher correlates of protective immunity. In light of possible Zika-Dengue bi-directional antibody-dependent enhancement (ADE) of disease (Dejnirattisai et al., 2016), availability of surrogate animal models of both viruses to predict safety and clinical benefit of candidate vaccines is essential. Cynomolgus macaques have been established as a model for Dengue virus infection and are currently be used for preclinical safety and efficacy of vaccine candidates. The objective of this study was to establish proof of concept Zika natural history data in a limited number of cynomolgus monkey's to justify continued model development including an ADE model. Taken together, our preliminary data supports continued development of cynomolgus macaques as a model for ZIKV infection and certainly warrant further investigation with more statistically relevant numbers of animals. A cynomolgus macaque model of ZIKV infection could be a useful tool to understand the ZIKV natural infection, pathology and develop effective control measures.

AUTHOR CONTRIBUTIONS

FK is the Principal Investigator and Study Director of this study. He designed the study, and authored the study and IACUC protocols. SG developed the RT-qPCR assay, analyzed the results and contributed to manuscript preparation. BS is a scientist who propagated Zika virus and performed Zika virus plaque assays. KW is a scientist who directed virus propagation and characterization, contributed to the protocols and manuscript preparation. AG is a scientist who performed all RNA extractions and contributed to RT-qPCR. KH conducted NHP handling, virus dosing of NHPs, sample collections and necropsy of animals. RK is a scientist who developed the *in vitro* protocols for this project, advised SG and BS and contributed to manuscript preparation. JR is a Co-PI and supported study design and study protocol development. He also supported proposal writing to secure funding for this work.

FUNDING

This work was performed with the internal funds from Southern Research.

ACKNOWLEDGMENTS

We would like to acknowledge the source of our ZIKV FSS13025 stock received from UTMB, Galveston National Laboratory, 301 University Boulevard, Galveston, TX 77550, USA; and ZIKV PRVABC59 from the CDC, 3156 Rampart Road, Fort Collins, CO 80526, USA.

REFERENCES

Abbink, P., Larocca, R. A., De La Barrera, R. A., Bricault, C. A., Moseley, E. T., Boyd, M., et al. (2016). Protective efficacy of multiple vaccine platforms against Zika virus challenge in rhesus monkeys. *Science* 353, 1129–1132.

Andrade, M. C., Ribeiro, C. T., Silva, V. F., Molinaro, E. M., Goncalves, M. A., Marques, M. A., et al. (2004). Biologic data of *Macaca mulatta*, *Macaca fascicularis*, and *Saimiri sciureus* used for research at the Fiocruz primate center. *Mem. Inst. Oswaldo Cruz* 99, 581–589.

Antony, J. M., and MacDonald, K. S. (2015). A critical analysis of the cynomolgus macaque, Macaca fascicularis, as a model to test HIV-1/SIV vaccine efficacy. *Vaccine* 33, 3073–3083. doi: 10.1016/j.vaccine.2014.12.004

Brasil, P., Calvet, G. A., Siqueira, A. M., Wakimoto, M., de Sequeira, P. C., Nobre, A., et al. (2016). Zika Virus Outbreak in Rio de Janeiro, Brazil: Clinical Characterization, Epidemiological and Virological Aspects. *PLoS Negl. Trop. Dis.* 10:e0004636. doi: 10.1371/journal.pntd.0004636

Brasil, P., and Nielsen-Saines, K. (2016). More pieces to the microcephaly-Zika virus puzzle in Brazil. *Lancet Infect. Dis.* doi: 10.1016/S1473-3099(16)30372-3 [Epub ahead of print].

Brault, J. B., Khou, C., Basset, J., Coquand, L., Fraisier, V., Frenkiel, M. P., et al. (2016). Comparative analysis between flaviviruses reveals specific neural stem cell tropism for Zika virus in the mouse developing neocortex. *EBioMedicine* 10, 71–76. doi: 10.1016/j.ebiom.2016.07.018

Calvet, G., Aguiar, R. S., Melo, A. S., Sampaio, S. A., de Filippis, I., Fabri, A., et al. (2016). Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect. Dis.* 16, 653–660. doi: 10.1016/S1473-3099(16)00095-5

Campos, G. S., Bandeira, A. C., and Sardi, S. I. (2015). Zika Virus Outbreak, Bahia, Brazil. *Emerg. Infect. Dis.* 21, 1885–1886. doi: 10.3201/eid2110.150847

Cox, J., Mota, J., Sukupolvi-Petty, S., Diamond, M. S., and Rico-Hesse, R. (2012). Mosquito bite delivery of dengue virus enhances immunogenicity and pathogenesis in humanized mice. *J. Virol.* 86, 7637–7649. doi: 10.1128/JVI.00534-12

Cugola, F. R., Fernandes, I. R., Russo, F. B., Freitas, B. C., Dias, J. L., Guimaraes, K. P., et al. (2016). The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* 534, 267–271.

Dejnirattisai, W., Supasa, P., Wongwiwat, W., Rouvinski, A., Barba-Spaeth, G., Duangchinda, T., et al. (2016). Dengue virus sero-cross-reactivity drives antibody-dependent enhancement of infection with zika virus. *Nat. Immunol.* 17, 1102–1108. doi: 10.1038/ni.3515

Dick, G. W., Kitchen, S. F., and Haddow, A. J. (1952). Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hygiene* 46, 509–520. doi: 10.1016/0035-9203(52)90042-4

do Rosario, M. S., de Jesus, P. A., Vasilakis, N., Farias, D. S., Novaes, M. A., Rodrigues, S. G., et al. (2016). Guillain-Barre Syndrome After Zika Virus Infection in Brazil. *Am. J. Trop. Med. Hygiene* 95, 1157–1160. doi: 10.4269/ajtmh.16-0306

Dowall, S. D., Graham, V. A., Rayner, E., Atkinson, B., Hall, G., Watson, R. J., et al. (2016). A Susceptible Mouse Model for Zika Virus Infection. *PLoS Negl. Trop. Dis.* 10:e0004658. doi: 10.1371/journal.pntd.0004658

Dudley, D. M., Aliota, M. T., Mohr, E. L., Weiler, A. M., Lehrer-Brey, G., Weisgrau, K. L., et al. (2016). A rhesus macaque model of Asian-lineage Zika virus infection. *Nat. Commun.* 7:12204. doi: 10.1038/ncomms12204

Duffy, M. R., Chen, T. H., Hancock, W. T., Powers, A. M., Kool, J. L., Lanciotti, R. S., et al. (2009). Zika virus outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* 360, 2536–2543. doi: 10.1056/NEJMoa0805715

Enfissi, A., Codrington, J., Roosblad, J., Kazanji, M., and Rousset, D. (2016). Zika virus genome from the Americas. *Lancet* 387, 227–228. doi: 10.1016/S0140-6736(16)00003-9

Goebel, S., Snyder, B., Sellati, T., Saeed, M., Ptak, R., Murray, M., et al. (2016). A sensitive virus yield assay for evaluation of Antivirals against Zika Virus. *J. Virol. Methods* 238, 13–20. doi: 10.1016/j.jviromet.2016.09.015

Gupta, A. K., Kaur, K., Rajput, A., Dhanda, S. K., Sehgal, M., Khan, M. S., et al. (2016). ZikaVR: an integrated Zika virus resource for genomics, proteomics, phylogenetic and therapeutic analysis. *Sci. Rep.* 6:32713. doi: 10.1038/srep32713

Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., et al. (2012). Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl. Trop. Dis.* 6:e1477. doi: 10.1371/journal.pntd.0001477

Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Velez, J. O., Lambert, A. J., Johnson, A. J., et al. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* 14, 1232–1239. doi: 10.3201/eid1408.080287

Lanciotti, R. S., Lambert, A. J., Holodniy, M., Saavedra, S., and Signor Ldel, C. (2016). Phylogeny of Zika Virus in Western Hemisphere, 2015. *Emerg. Infect. Dis.* 22, 933–935.

Lazear, H. M., Govero, J., Smith, A. M., Platt, D. J., Fernandez, E., Miner, J. J., et al. (2016). A Mouse Model of Zika Virus Pathogenesis. *Cell Host Microbe* 19, 720–730. doi: 10.1016/j.chom.2016.03.010

Martines, R. B., Bhatnagar, J., de Oliveira Ramos, A. M., Davi, H. P., Iglezias, S. D., Kanamura, C. T., et al. (2016). Pathology of congenital Zika syndrome in Brazil: a case series. *Lancet* 388, 898–904. doi: 10.1016/S0140-6736(16)30883-2

Mlakar, J., Korva, M., Tul, N., Popovic, M., Poljsak-Prijatelj, M., Mraz, J., et al. (2016). Zika virus associated with microcephaly. *N. Engl. J. Med.* 374, 951–958. doi: 10.1056/NEJMoa1600651

Musso, D., Nilles, E. J., and Cao-Lormeau, V. M. (2014). Rapid spread of emerging Zika virus in the Pacific area. *Clin. Microbiol. Infect.* 20, O595–O596. doi: 10.1111/1469-0691.12707

Oehler, E., Watrin, L., Larre, P., Leparc-Goffart, I., Lastere, S., Valour, F., et al. (2014). Zika virus infection complicated by Guillain-Barre syndrome–case report, French Polynesia, December 2013. *Euro Surveill.* 19:20720. doi: 10.2807/1560-7917.ES2014.19.9.20720

Osuna, C. E., Lim, S. Y., Deleage, C., Griffin, B. D., Stein, D., Schroeder, L. T., et al. (2016). Zika viral dynamics and shedding in rhesus and cynomolgus macaques. *Nat. Med.* 22, 1448–1455. doi: 10.1038/nm.4206

Plourde, A. R., and Bloch, E. M. (2016). A Literature Review of Zika Virus. *Emerg. Infect. Dis.* 22, 1185–1192. doi: 10.3201/eid2207.151990

Sarno, M., Aquino, M., Pimentel, K., Cabral, R., Costa, G., Bastos, F., et al. (2016). Progressive lesions of Central Nervous System in microcephalic fetuses with suspected congenital Zika virus syndrome. *Ultrasound obstet. Gynecol.* doi: 10.1002/uog.17303 [Epub ahead of print].

Solomon, T., Baylis, M., and Brown, D. (2016). Zika virus and neurological disease–approaches to the unknown. *Lancet Infect. Dis.* 16, 402–404. doi: 10.1016/S1473-3099(16)00125-0

Spearman, P. (2006). Current progress in the development of HIV vaccines. *Curr. Pharm. Des.* 12, 1147–1167. doi: 10.2174/138161206776055859

Styer, L. M., Kent, K. A., Albright, R. G., Bennett, C. J., Kramer, L. D., and Bernard, K. A. (2007). Mosquitoes inoculate high doses of West Nile virus as they probe and feed on live hosts. *PLoS Pathog.* 3:1262–1270.

van Hemert, F., and Berkhout, B. (2016). Nucleotide composition of the Zika virus RNA genome and its codon usage. *Virol. J.* 13:95. doi: 10.1186/s12985-016-0551-1

Wang, L., Valderramos, S. G., Wu, A., Ouyang, S., Li, C., Brasil, P., et al. (2016). From Mosquitos to humans: genetic evolution of Zika Virus. *Cell Host Microbe* 19, 561–565. doi: 10.1016/j.chom.2016.04.006

Ye, Q., Liu, Z. Y., Han, J. F., Jiang, T., Li, X. F., and Qin, C. F. (2016). Genomic characterization and phylogenetic analysis of Zika virus circulating in the Americas. *Infect. Genet. Evol.* 43, 43–49. doi: 10.1016/j.meegid.2016.05.004

Zanluca, C., Melo, V. C., Mosimann, A. L., Santos, G. I., Santos, C. N., and Luz, K. (2015). First report of autochthonous transmission of Zika virus in Brazil. *Mem. Inst. Oswaldo Cruz* 110, 569–572. doi: 10.1590/0074-02760150192

Zmurko, J., Marques, R. E., Schols, D., Verbeken, E., Kaptein, S. J., and Neyts, J. (2016). The viral polymerase inhibitor 7-Deaza-2′-C-Methyladenosine is a potent inhibitor of in vitro Zika virus replication and delays disease progression in a robust mouse infection model. *PLoS Negl. Trop. Dis.* 10:e0004695. doi: 10.1371/journal.pntd.0004695

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Koide, Goebel, Snyder, Walters, Gast, Hagelin, Kalkeri and Rayner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 November 2016
doi: 10.3389/fmicb.2016.01688

Systems Biomedicine of Rabies Delineates the Affected Signaling Pathways

Sadegh Azimzadeh Jamalkandi[1†], Sayed-Hamidreza Mozhgani[2†], Hamid Gholami Pourbadie[3], Mehdi Mirzaie[4], Farshid Noorbakhsh[5], Behrouz Vaziri[6], Alireza Gholami[7], Naser Ansari-Pour[8,9]* and Mohieddin Jafari[10]**

[1] Chemical Injuries Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran, [2] Department of Virology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran, [3] Department of Physiology and Pharmacology, Pasteur Institute of Iran, Tehran, Iran, [4] Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran, [5] Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran, [6] Protein Chemistry and Proteomics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran, [7] WHO Collaborating Center for Reference and Research on Rabies, Pasteur Institute of Iran, Tehran, Iran, [8] Faculty of New Sciences and Technology, University of Tehran, Tehran, Iran, [9] Department of Genetics, Evolution and Environment, UCL Genetics Institute, University College London, London, UK, [10] Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

The prototypical neurotropic virus, rabies, is a member of the Rhabdoviridae family that causes lethal encephalomyelitis. Although there have been a plethora of studies investigating the etiological mechanism of the rabies virus and many precautionary methods have been implemented to avert the disease outbreak over the last century, the disease has surprisingly no definite remedy at its late stages. The psychological symptoms and the underlying etiology, as well as the rare survival rate from rabies encephalitis, has still remained a mystery. We, therefore, undertook a systems biomedicine approach to identify the network of gene products implicated in rabies. This was done by meta-analyzing whole-transcriptome microarray datasets of the CNS infected by strain CVS-11, and integrating them with interactome data using computational and statistical methods. We first determined the differentially expressed genes (DEGs) in each study and horizontally integrated the results at the mRNA and microRNA levels separately. A total of 61 seed genes involved in signal propagation system were obtained by means of unifying mRNA and microRNA detected integrated DEGs. We then reconstructed a refined protein–protein interaction network (PPIN) of infected cells to elucidate the rabies-implicated signal transduction network (RISN). To validate our findings, we confirmed differential expression of randomly selected genes in the network using Real-time PCR. In conclusion, the identification of seed genes and their network neighborhood within the refined PPIN can be useful for demonstrating signaling pathways including interferon circumvent, toward proliferation and survival, and neuropathological clue, explaining the intricate underlying molecular neuropathology of rabies infection and thus rendered a molecular framework for predicting potential drug targets.

Keywords: rabies, systems biology, protein–protein interaction network, signaling network, microarray, real-time PCR

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Takashi Irie,
Hiroshima University, Japan
Iman Tavassoly,
Icahn School of Medicine at Mount Sinai, USA

***Correspondence:**
Alireza Gholami
a.gholami@pasteur.ac.ir
Naser Ansari-Pour
n.ansaripour@ut.ac.ir
Mohieddin Jafari
m_jafari@pasteur.ac.ir;
mjafari@ipm.ir;
www.jafarilab-pasteur.com

[†]These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *10 August 2016*
Accepted: *07 October 2016*
Published: *07 November 2016*

Citation:
Azimzadeh Jamalkandi S, Mozhgani S-H, Gholami Pourbadie H, Mirzaie M, Noorbakhsh F, Vaziri B, Gholami A, Ansari-Pour N and Jafari M (2016) Systems Biomedicine of Rabies Delineates the Affected Signaling Pathways. Front. Microbiol. 7:1688. doi: 10.3389/fmicb.2016.01688

INTRODUCTION

Growing evidence of inter-population and inter-individual variation in the attack rate and prognosis of specific infectious diseases suggest an underlying biological complexity. In fact, any perturbation in the densely organized inter-relationship of genetic and environmental factors may lead to this intricate behavior (Hunter, 2005). In particular, the strange survival pattern observed from fatal rabies infection of the central nervous system (CNS) introduces this infection as a complex disease (de Souza and Madhusudana, 2014).

The prototypical neurotropic virus, rabies, is a member of the Rhabdoviridae family that causes lethal encephalomyelitis (Sugiura et al., 2011). The viruses in this family are enveloped with a single stranded negative sense RNA genome. The genomic length of the rabies virus (RABV) is about 12 kb and encodes five proteins including nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and a viral RNA polymerase (L) (Yousaf et al., 2012). This neglected virus leads to death once the symptoms develop and has a mortality rate of 1:100,000 to 1:1000 per year. The deceased intriguingly display no neural damage, neurohistopathological evidence, or induced severe immune response (Schnell et al., 2010). In an organized hijacking program, the virus travels from the muscle tissue to the nervous system, migrates to the spinal cord and freely covers certain parts of the brain (Schnell et al., 2010). The virus spreads centrifugally to other organs and subsequently to the next host. Although the host innate immune response including TLR, type 1 interferon, TNF alpha, and IL-6 are the first defense line against a viral infection, this virus easily propagates in the nervous system. This suggests that the RABV has a specific mechanism to suppress host innate immunity (Rupprecht, 1996; Ito et al., 2011; Gomme et al., 2012). Several laboratory strains of the RABV in addition to the wild types cause fatal acute encephalomyelitis associated with inflammation of the brain and spinal cord, leading to coma and death especially when the virus is injected intracerebrally in high dose (Meslin et al., 1996; Galelli et al., 2000; Baloul and Lafon, 2003; Baloul et al., 2004). In contrast to attenuated strains, wild type strains and CVS-11 do not induce histopathological changes indicative of apoptosis or necrosis in infected cells (Thoulouze et al., 1997, 2003a,b; Lay et al., 2003; Préhaud et al., 2003). Accordingly, despite over 100 years of controlling rabies by developing RABV vaccines and serotherapy, the precise neurological and immunological etiology as well as rare survival cases from rabies encephalitis still remains a mystery (Gomme et al., 2012; de Souza and Madhusudana, 2014).

After the emergence of omics technology, some studies have started to pave the way toward a better understanding of rabies fatal mechanism. Elucidating the essential biological processes involved in rabies progression has been based mainly on analyzing gene expression alterations. Zhao et al. reported expression profiling of mRNA and microRNA of rabies-infected cell (Zhao et al., 2011, 2012a,b, 2013). Suigiura et al. analyzed the gene expression profile of CNS tissue infected with CVS-11 (Sugiura et al., 2011). Changes in gene expression were also studied in marked neurons infected with recombinant RABV expressing CRE-recombinase (Gomme et al., 2012). Numerous other studies have also analyzed gene expression profiling using transcriptomic or proteomic methods within diverse cellular models in different species (Wang et al., 2005, 2011; Dhingra et al., 2007; Fu et al., 2008; Zandi et al., 2009, 2013; Han et al., 2011; Thanomsridetchai et al., 2011; Vaziri et al., 2012; Farahtaj et al., 2013; Francischetti et al., 2013; Kluge et al., 2013; Silva et al., 2013; Venugopal et al., 2013; Kasempimolporn et al., 2014; Kammouni et al., 2015; Mehta et al., 2015).

To increase the reliability of results and generalizability of these independent but related studies, it is recommended to statistically combine such data, commonly known as data integration or meta-analysis (Ramasamy et al., 2008). Several studies have shown the benefits of meta-analysis in terms of both higher statistical power and precision in detecting differentially expressed genes (DEGs) in different complex traits including infectious disease (Song et al., 2014; Camacho-Cáceres et al., 2015; Sharma et al., 2015; Yin et al., 2015; Wang C.-Y. et al., 2015; Wang X. et al., 2015). Further, data integration approaches at a higher level try to map multiple biological data levels into one mechanistic network to improve representativeness of data (Chen et al., 2008; Bowick and McAuley, 2011; Amiri et al., 2013; Depiereux et al., 2015; Paraboschi et al., 2015). The generated multi-dimensional network is likely to be more useful in inferring universally involved processes or pathways regardless of inter-studies differences (Azimzadeh Jamalkandi et al., 2015).

Having in mind the common concerns in meta-analysis, we horizontally integrated nine high-throughput transcriptome datasets to identify consensus DEGs. The underlying molecular network in rabies pathogenesis was then extracted based on protein–protein interaction network (PPIN) and signaling pathways by defining the identified DEGs as seed genes. Finally, using real-time PCR, we experimentally validated a number of key DEGs in rabies-infected cells. We demonstrate that a systems biomedicine approach, based on integrating omics datasets and experimental validation, may be used to shed light on a vague portrait of a complex disease pathobiology.

METHODS

Super Horizontal Integration

Datasets

We looked into all databases pertaining to microarray data at both levels of mRNA and microRNA. This was done by searching databases [Gene Expression Omnibus (GEO), ArrayExpress, Google Scholar, and PubMed NCBI] and studies regarding the rabies virus were extracted with rabies-related keywords including "rabies", "RABV," and "rhabdoviridae" (**Figure 1A**). Out of a total of 13 studies, 9 were selected for further analysis. The list of included studies and their respective features are given in **Supplementary Table 1**. In the majority of studies, "brain" and "brain spinal" were the tissues under investigation, and the rest were examined on *Mus musculus*-derived microglial cells. It should also be noted that only two out of nine inclusive studies were conducted on both mRNA and microRNA levels, with others analyzing only one level of data. Four, three, and two of

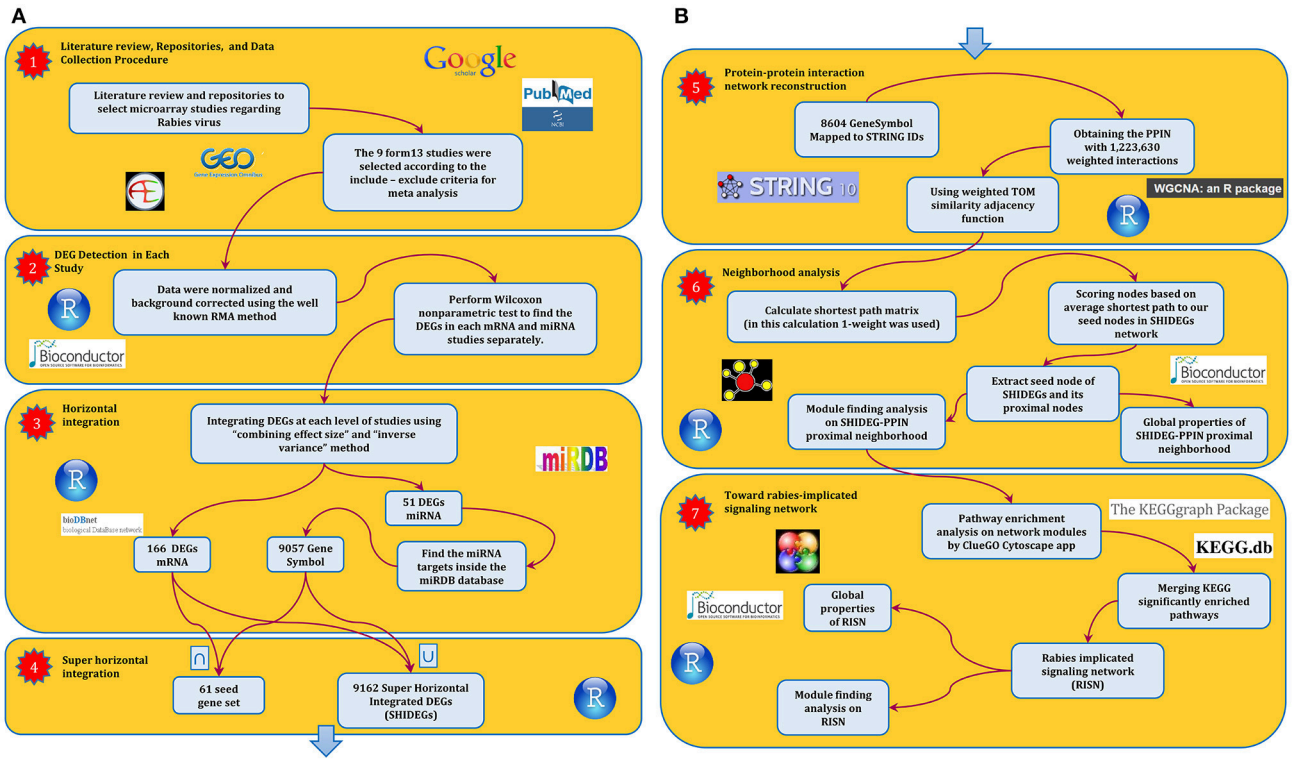


FIGURE 1 | The abstract flowchart of this study design. (A) The four steps taken to obtain the primitive interaction network of the RABV infection. The first step was a systematic review of the literature regarding the RABV. GEO was used to obtain the expression values of studies at the two levels of expression profiling by array of coding and non-coding RNA. Having selected the studies according to our data integration criteria, we moved on to the second step in which we detected DEGs at each level using non-parametric methods. The third step included integration of results using the meta-analysis techniques described in Ramasamy et al. (2008). The implemented integration method revealed that 166 mRNA and 51 microRNA (9057 microRNA targets) are differentially expressed. The next step consisted of super horizontal integration of all transcriptome data. Finally, the 9162 expressed genes were mapped to STRING v 10.0 for further analysis. **(B)** The next stage of the approach which comprised three steps resulted in the rabies-implicated signaling network (RISN). Firstly, the PPIN of all 9162 genes was reconstructed using STRING v10. The combined score calculated in STRING was used as edge weight in the SHIDEG-PPIN. In the second step, the proximal nodes of the seed gene set (nodes) in the whole SHIDEG-PPIN were found to create the seed neighborhood network. Module finding was undertaken along with the global network analysis. Next, the significantly enriched signaling pathways were extracted in the network modules separately. These pathways were then merged together, forming the rabies-implicated signaling network (RISN). Finally, analysis the global network and functional module finding analysis was performed followed by biological inference.

these studies were performed on samples infected by CVS-11, FJDRV and ERA, and RABV-Cre, respectively.

Data Normalization

In order to prepare data for integration and detect DEGs, it is necessary to use preprocessed and background corrected microarray-data (Ramasamy et al., 2008). First, we checked the quality of recorded CEL format data all of which required to be normalized. The data were normalized by using the "Affy" package in Bioconductor (Gautier et al., 2004). This includes between and within array normalization which reduces the effect of noise and contributes to data consistency. The MA plot and qq-plot for the pairs of samples in each study was analyzed separately to check the normality of data after normalization as a quality control step. Although the qq-plot of some studies revealed that the normalization methods had worked fine, normalization was not successful in datasets with significantly small sample sizes mainly because normality assumptions are violated in low sample size studies.

Analysis of Differentially Expressed Genes (DEGs)

Assuming that microarray data are normally distributed, the routine procedure to detect DEGs is to perform *t*-test, however, this may result in misleading conclusions if the normality assumption is violated. A recent study showed that oligonucleotide expression values, resulting from widely acceptable calculation methods, are not normally distributed (Hardin and Wilson, 2009). This suggests that the results of *t*-test are biased and unreliable, especially when the sample size in each group is significantly small, and more robust methods should be implemented. Here, we used the Wilcoxon–Mann–Whitney non-parametric test as an alternative method to identify DEGs with the significance level set to 0.05. The next step was to remove the un-mapped probes and solve the problem of "many-to-many conversion" as described in Ramasamy et al. (2008). This was done for both studies of mRNA and microRNA. The computational scripts plus an example of raw data are provided in **Data Sheet 1**.

Meta-Analysis

The results at each level (mRNA and microRNA) were integrated separately, using the inverse-variance technique and combining effect sizes as described in Ramasamy et al. (2008). For integration purposes, the list of DEGs in each study was gathered and the effect size value of each gene was then calculated. We only selected genes with an absolute effect size value >0.8 or those with a fold-change <0.33 or >1.5 (in at least one study) as the frequently accepted cut-off for fold-change. We obtained two different lists of significant differentially expressed values for mRNA and microRNA, respectively. We obtained the target gene symbols of each microRNA accession ID using miRDB (http://mirdb.org; Wong and Wang, 2015). To be best of our knowledge, for *Mus musculus*, this is the most up-to-date repository to convert accession IDs. Finally, after these two parallel horizontal integrations, the union and intersection of the results were extracted as super-horizontally integrated DEGs (SHIDEGs) and the seed gene set of SHIDEGs, respectively.

Background Network Construction

We used the STRING database to constructed a large-scale PPIN from seven available interaction sources and chose the lowest cut-off for combined scores (Downloaded on 2 September 2015; Szklarczyk et al., 2014). A total of 8604 proteins based on 9162 SHIDEGs were represented in STRING. Accordingly, a total of 1,223,630 edges were extracted and the STRING combined scores were used as edge weights. Next, the weighted adjacency matrix was transformed to a new adjacency matrix using topological overlapping measure (TOM) function in WGCNA package of R software (Yip and Horvath, 2007; Langfelder and Horvath, 2008; Song et al., 2012). It should be noted that the TOM transformation increases the non-zero adjacency matrix elements as well as very low weight values in this case. The transformed weight distributions of STRING default cut-offs, from lowest to highest confidence, were thus considered to define a new threshold. The third quartile of transformed scores (0.4577) of the highest confidence was selected to strictly filter weak and false-positive interactions.

Neighborhood Ranking

Using the custom igraph package in R, we generated a matrix of all shortest paths between all pairs of nodes in a weighted network with the algorithm of Dijkstra (Csardi and Nepusz, 2006). First, we substituted raw weights with one-weight to increase reachability of nodes with high weights to seed gene set (nodes) in the shortest path finding procedure. We then defined a distance score, D_j, for each node in the PPIN as the difference in average of the shortest path to the node when starting on a non-seed node compared with when starting on a seed node, normalized by the average shortest path to reach the node from the whole network.

$$D_j = \frac{\frac{\sum_{i \notin S} SP_{ij}}{|NS|} - \frac{\sum_{i \in S} SP_{ij}}{|S|}}{\frac{\sum_i SP_{ij}}{|S| + |NS|}}$$

Here S is the set of nodes that fall into the seed gene set and NS is the set of nodes that are non-seed nodes. Therefore, a score greater than zero implies that node j falls closer on average to the seed nodes than it does on average to the rest of the network. The rabies network was generated based on the SHIDEGs seed gene set and each member of the seed gene set by scoring all nodes in the network and using a cutoff score of zero to define the neighborhood. It should be noted that the D scores were calculated without imposing any threshold on edge weights.

Undirected PPIN; Topological and Pathway Enrichment Analysis

To reconstruct a high confidence PPIN around our seed gene set, we used the 0.4577 threshold to filter weak interaction among neighborhood nodes. This filtering resulted in the proximal neighborhood network of seeds. Using Gephi version 0.9, the global topological properties of the resulting PPIN along with module identification was analyzed. To undertake enrichment analysis among the detected modules, ClueGO 2.1.7 (Bindea et al., 2009) in Cytoscape 3.2.1 was used based on *Mus musculus* using the following parameters: KEGG (Kanehisa et al., 2014), Reactome (Croft et al., 2014), and Wikipathway ontology databases (Kelder et al., 2012), default term selection options, hypergeometric test and Bonferroni step-down p-value correction.

Signaling Network Analysis

The rabies-implicated signaling network (RISN) was constructed based on the KEGG pathways enriched in the rabies PPIN. All statistically significant and frequent pathways in all PPIN modules were extracted and merged together to build a large-scale RISN. All SHIDEGs were then delineated in this network by different color labeling. After reviewing clinical and physiological evidences pertaining to the RABV, the whole RISN was delineated into a less complex network.

Cell Culture and Virus

The Neuro-2a cell line, a murine neuroblastoma cell line, and CVS-11 strain of the RABV (the challenge virus standard) were obtained from the WHO collaborating center for reference and research on rabies, Pasteur Institute of Iran (Tehran, Iran). Virus titers were determined by a focal infectivity assay using BSR (a line of BHK) cells. Neuro-2a cells were grown in Dulbecco's Modified Eagle Medium (DMEM) containing 4500 mg/L glucose and sodium bicarbonate, supplemented with 10% fetal bovine serum. Cultures were maintained at 37 C in a 5% CO_2 humidified cell incubator with growth medium replaced every 48 h. For all experiments, cells were subcultured into 25 cm tissue culture flasks and were grown for 16 h before infection.

Total RNA Isolation, cDNA Synthesis, and Primer Design for PCR

Total mRNA was isolated from neuroblastoma cells (mock infected and infected with the CVS-11 strain of RABV) using the RNX RNA Isolation Kit (CinnaGen Inc., Tehran, Iran). The amount and purity of RNA were determined by Biotek microplate spectrophotometry. The extracted RNA was then treated with DNase to remove genomic DNA.

Total RNA (1.7 μg/ml) was reverse transcribed into first-strand cDNA by the SuperScript III First-Strand Synthesis System (Thermo Fisher Scientific) and oligo(dT)18 according to the manufacturer's protocol. Primer specificity was tested by primer-BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and experimentally by the positive control amplification. Optimal PCR conditions were identified for each primer pair. GAPDH was used as an internal control for RT-polymerase chain reactions.

Quantitative Real-Time PCR

Reverse transcription-quantitative real-time PCR (RT-qPCR) was carried out on a Rotor-Gene Q 5plex HRM instrument (Qiagen, Hilden, Germany) with EvaGreen fluorescence dye (Biotium, Hayward, USA) to monitor cDNA amplification of *GNAI2*, *AKT3*, *IL21*, and *GAPDH* through increased fluorescence intensity. The specificity of the amplified products was checked by melting curve analysis, and the expected size of the fragments was further visualized by gel electrophoresis (2% agarose) and staining with GelRed (Biotium, Hayward, CA). Results were confirmed by triplicate testing. Relative mRNA expression was calculated using the delta–delta *Ct* method (Livak and Schmittgen, 2001). Sequences were analyzed using Seqscanner. Statistical analysis was performed by depicting an error bar for each gene in each condition to compare relative expression of the abovementioned genes in uninfected and RABV-infected states.

RESULTS

This study comprises seven steps in two separate parts as illustrated in the **Figure 1**. After a systematic literature review, nine transcriptomic datasets pertaining to rabies were collected (**Supplementary Table 1**). DEGs were identified in the mRNA and microRNA datasets and used to construct a PPIN of rabies infection. In the second part, analysis of the PPIN neighborhood and rabies-implicated signaling was implemented to create a mechanistic description of the molecular pathogenesis of rabies infection (**Figure 1B**). The outcome of each step is discussed in more detail below.

Intersection of mRNA and microRNA Transcriptome Data by Super Horizontal Integration Reveals an Intriguing List of the Seed Gene Set

A total of 166 DEGs were identified at the mRNA level. Analysis at the microRNA level led to the identification of 51 genes which target 9057 genes on mouse genome. The genes at both mRNA and microRNA levels were combined to create a list of 9162 "super horizontally integrated-DEGs (SHIDEGs)" and a list of 61 intersecting genes considered as "seed genes" (**Table 1**). Of the 9162 SHIDEGs, 8604 (∼93%) were mapped to STRING version 10 and the obtained network contained 1,223,630 weighted protein–protein interactions.

Next, we obtained gene ontology (GO) classifications for all seed genes. Using the Enrichr web based tools (Chen et al., 2013; Kuleshov et al., 2016) significantly enriched biological process (BP), molecular function (MF), and cellular component (CC) terms were retrieved and then ranked by combined scores (**Figure 2**). From a biological process point of view, diverse inflammatory responses such as cytokine-mediated signaling pathway (GO:0019221) and regulation of leukocyte activation (GO:0002694) were enriched. Also, the other moiety of BPs was generally associated with nucleotide biosynthetic processes (**Figure 2A**). This observation confirmed the role of immune signaling pathways and propagation apparatus in rabies infection. CC and MF enriched terms further accentuated the role of signaling alteration in development of rabies (**Figures 2B,C**).

Shortest Path-Based Scoring Allows Identification of the Seed Gene Neighborhood in the Protein–Protein Interaction Space

We applied network concepts to explore more thoroughly the potential functional relationship between the identified DEGs and RABV pathogenesis. We postulated that all integrated DEGs are involved in the global interactome perturbed by the RABV. We assumed that the SHIDEGs are more likely to interact directly with the RABV and the neighbors of SHIDEGs are of the next level of etiological importance. To identify the disease subnetwork of SHIDEGs in PPIN, we retrieved the entire protein–protein weighted interactions from STRING (Szklarczyk et al., 2014). The giant component comprising 8604 DEG products was selected for further analysis. Given the high false-positive rate in PPINs (Jafari et al., 2013, 2015), the topological overlap matrix (TOM)-based adjacency function was used to filter the effect of spurious or weak connections (Li and Horvath, 2007; Yip and Horvath, 2007). Proteins encoded by the 61 seed genes were identified in the refined global PPIN for neighborhood analysis.

Based on biological parsimony and the observed patterns in different signaling databases, biological responses are controlled via a short signaling cascade (Gitter et al., 2011; Silverbush and Sharan, 2014). We therefore used the shortest path algorithm to identify nodes in proximity of the seed nodes. We then ranked the nodes within the whole PPIN using distance D. From the total 8602 nodes within the robust PPIN, 3775 nodes had a positive score and therefore fell within the seed gene neighborhood.

Subsequently, to filter the edges having low weight, STRING combined score (weight of edges) were transformed using the TOM-based adjacency function and those above the 3rd quartile were retained. Of all nodes with $D > 0$, 694 nodes passed this filter and were considered as the "seed gene proximal neighborhood network". This resulted in the selection of highly important relationships among nodes based on network topology (**Figure 3**).

The Identification of Rabies-Implicated Gene Products

The final rabies infection PPIN contained 694 nodes with 6097 interactions. The degree distribution (**Supplementary Figure 1**) and modularity index (∼0.7) of this PPIN indicate that it has a modular structure and a scale free topology. Its average path

TABLE 1 | Seed gene set.

Synonyms	UniProt accession number	Name	mRNA expression	miRNA expression
Myd88	P22366	Myeloid differentiation primary response 88	Up	Down
Il7r	P16872	Interleukin 7 receptor	Up	Down
Ccnb1	P24860	Cyclin B1	Up	Down
B4galt1	P15535	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 1	Up	Down
Cmya5	Q70KF4	Cardiomyopathy associated 5	Down	Up
Amotl1	Q9D4H4	Angiomotin like 1	Down	Up
Styk1	Q6J9G1	Serine/Threonine/Tyrosine kinase 1	Down	Up
HAP1	O35668	Huntingtin-associated protein 1	Down	Up
Slc16a4	Q8R0M8	Solute carrier family 16, member 4	Down	Up
Cyth4	Q80YW0	Cytohesin 4	Down	Up
Ctsz	Q9WUU7	Cathepsin Z	Down	Up
Pnp	P23492	Purine nucleoside phosphorylase	Down	Up
Grap2	O89100	GRB2-related adaptor protein 2	Down	Up
Crem	P27699	CAMP responsive element modulator	Down	Up
CXCL10	P17515	Chemokine (C-X-C motif) ligand 10	Down	Up
Angptl4	Q9Z1P8	Angiopoietin-like 4	Down	Up
Baz1a	O88379	Bromodomain adjacent to zinc finger domain, 1A	Down	Up
KLRA3	Q64329	Killer cell lectin-like receptor 3	Down	Up
GBP2	Q9Z0E6	Guanylate binding protein 2, interferon-inducible	Down	Up
Lhx2	Q9Z0S2	LIM homeobox 2	Down	Up
Apod	P51910	Apolipoprotein D	Down	Up
Arhgap9	Q1HDU4	Rho GTPase activating protein 9	Down	Up
Gas1	Q01721	Growth arrest-specific 1	Down	Up
MITD1	Q8VDV8	MIT, microtubule interacting and transport, domain containing 1	Down	Up
Tom1l1	Q923U0	Target of myb1 (chicken)-like 1	Down	Up
Glipr2	Q9CYL5	GLI pathogenesis-related 2	Down	Up
Sh2d1b1	O35324	SH2 domain-containing protein 1B	Down	Up
A630001G21Rik	Q3UTB2	Protein A630001G21Rik	Down	Up
Ncr1	Q8C567	Natural cytotoxicity triggering receptor 1	Down	Up
Saa2	P05367	Serum amyloid A2	Down	Up
Ing5	Q9D8Y8	Inhibitor of growth family, member 5	Down	Up
Laptm5	Q61168	Lysosomal protein transmembrane 5	Down	Up
Kmo	Q91WN4	Kynurenine 3-monooxygenase (kynurenine 3-hydroxylase)	Down	Up
Dusp2	Q05922	Dual specificity phosphatase 2	Down	Up
Ctla2b	P12400	Protein CTLA-2-beta	Down	Up
Cd93	O89103	CD93 molecule	Down	Up
Serpinb1c	Q5SV42	Leukocyte elastase inhibitor C	Down	Up
4930486L24Rik	Q80UB0	Testin-2	Down	Up
USP18	Q9WTV6	Ubiquitin specific peptidase 18	Down	Up
Gosr1	O88630	Golgi SNAP receptor complex member 1	Down	Up
Tnfaip3	Q60769	Tumor necrosis factor, alpha-induced protein 3	Down	Up
IL15RA	Q60819	Interleukin 15 receptor, alpha	Down	Up
Eif4ebp1	Q60876	Eukaryotic translation initiation factor 4E binding protein 1	Down	Up
Fbxl5	Q8C2S5	F-box and leucine-rich repeat protein 5	Down	Up
Tnfrsf11a	O35305	Tumor necrosis factor receptor superfamily, member 11a, NFKB activator	Down	Down
Ly6c1	P0CW02	Lymphocyte antigen 6C1	Down	Down
Il1a	P01582	Interleukin 1, alpha	Down	Down
Aida	Q8C4Q6	Axin interactor, dorsalization associated	Down	Down
Mmp19	Q9JHI0	Matrix metallopeptidase 19	Down	Down

(Continued)

TABLE 1 | Continued

Synonyms	UniProt accession number	Name	mRNA expression	miRNA expression
Csf1	P07141	Colony stimulating factor 1 (macrophage)	Down	Down
Arrdc4	Q0GJK1	Arrestin domain containing 4	Down	Down
Nampt	Q99KQ4	Nicotinamide phosphoribosyltransferase	Down	Down
Nfkb2	Q9WTK5	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	Down	Down
Fkbp5	Q64378	FK506 binding protein 5	Down	Down
Klhl5	Q6PFE1	Kelch-like family member 5	Up	Up
Snx10	Q9CWT3	Sorting nexin 10	Up	Up
SERPINB9	O08797	Serpin peptidase inhibitor, clade B (ovalbumin), member 9	Up	Up
IGF2	P09535	Insulin-like growth factor 2	Up	Up
IFI204	P15092	Interferon-activable protein 204	Up	Up
Il17ra	Q60943	Interleukin 17 receptor A	Up	Up
Zc3h12a	Q5D1E7	Zinc finger CCCH-type containing 12A	Up	Up

The 61 seed genes are presented in this table. The green, red, and yellow color indicate over-, under, and ambivalent expression of genes regarding mRNA and microRNA expression evidence.

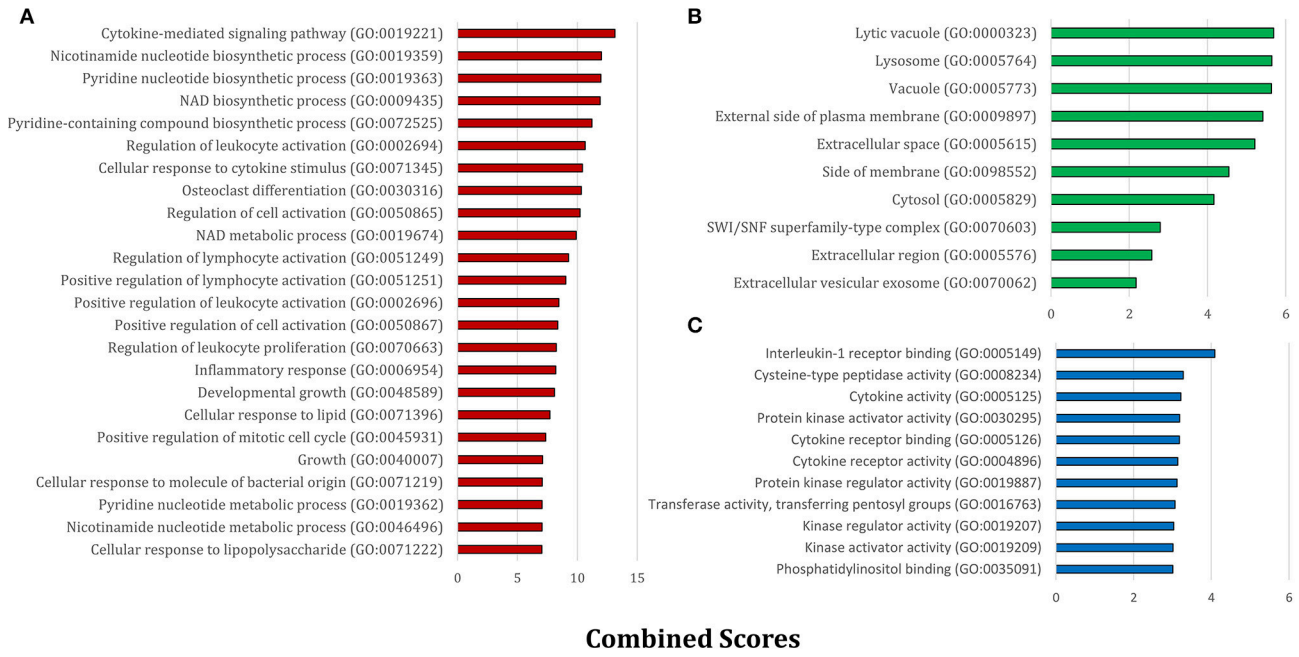


FIGURE 2 | The significant GO terms based on the seed gene set. The enriched GO terms within **(A)** biological processes, **(B)** cellular components, and **(C)** molecular functions are presented separately. All terms were statistically significant ($P < 0.05$) and are ranked based on Enrichr combined scores.

length and diameter were 5.16 and 14, respectively, showing that this relatively large and sparse network is small-world. To infer the functionality of this refined network, we analyzed the network modules. Twelve modules were detected by the fast unfolding clustering algorithm implemented in Gephi (V. 0.9) (Bastian et al.).

To avoid bias in inferring global properties of the network, the top three central nodes in each module were specifically shown in **Figure 4**. This Figure illustrates these nodes in terms of degree and betweenness centrality measures. Interestingly, all of these nodes have diverse receptor binding and kinase activity functions based on GO enrichment analysis. On the other hand, our results revealed that inter-modular high-degree nodes related to CCR chemokine receptor binding (GO:0048020), R-SMAD binding (GO:0070412), responses to mechanical stimulus (GO:0009612), JAK-STAT cascade involved in growth hormone signaling pathway (GO:0060397), and negative regulation of neuron death (GO:1901215) were down-regulated by the RABV. In contrast, the local and global hub proteins associated with positive regulation of protein kinase activity (GO:0045860),

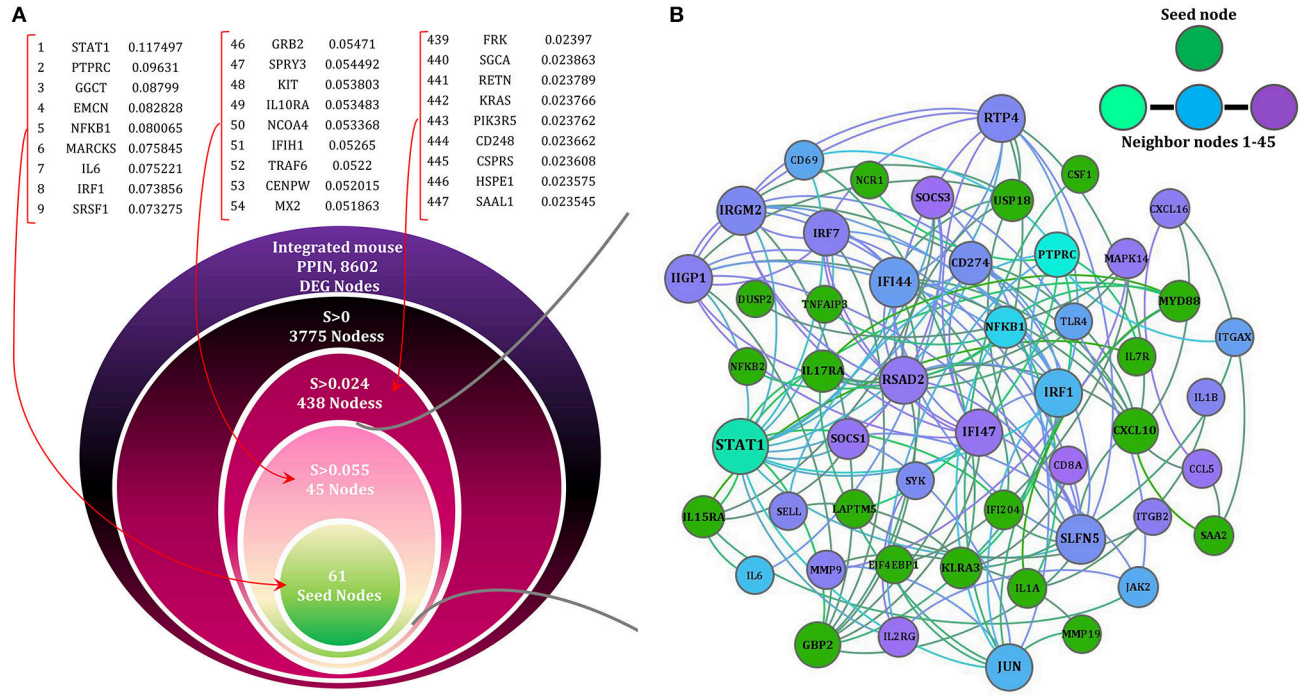


FIGURE 3 | The network (SHIDEG-PPIN) onion diagram. (A) Identification of the rabies disease neighborhood network based on proximal nodes of the seed gene set. The PPIN of all DEG products (8602/9162) existed in STRING were ranked on the basis of their shortest path (SP) score with the 61 rabies SHIDEGs as seed nodes. Selecting various score cutoffs (99.5th, 95th and 56th quantiles) allow neighborhoods of various sizes to be defined as shown in the nested circles. The ranked list of neighbors of seed nodes demonstrated with their scores. **(B)** The minimized PPIN formed with the confident interactions (0.4 cut-off selected based on the TOM procedure) between the 61 seed nodes (green) and top 45 ranked proximal nodes (range of colors from cyan to purple) is shown at the top with node size representing degree of nodes.

cellular response to lipid (GO:0071396), neurotrophin TRK receptor signaling pathway (GO:0048011), G-protein coupled receptor binding (GO:0001664), and neuropeptide hormone activity (GO:0005184) were up-regulated, thus facilitating virus survival and propagation by avoiding programmed cell death.

The same scenario also applied to nodes with high betweenness centrality. For example, nodes associated with neurotrophin receptor binding (GO:0005165) and cellular response to organonitrogen compound (GO:0071417) were over-expressed while those associated with natural immune system were underexpressed (**Supplementary Table 2**). On top of that, the top five ranked nodes based on betweenness centrality, namely EP300, STAT1, RHOA, and PDGFA were under-expressed concurrently. This may lead to a lack of network coordination among different immune processes.

Given the abundance of receptors and kinases in this network, we performed pathway enrichment analysis on each module separately. Using ClueGO (Cytoscape plugin; Bindea et al., 2009), the statistically significant pathway terms were identified among those in the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa et al., 2014) and Reactome (Croft et al., 2014) databases (**Table 2** and **Supplementary Table 3**). We then ranked the enriched pathway terms based on gene coverage (Ansari-Pour et al., 2016).

Furthermore, to evaluate the quality of module discovery results, conformity of enriched pathways in a module was assessed with respect to the interconnectedness level of that module (**Supplementary Table 4**). Our results demonstrated that the KEGG enriched pathway similarity matrix was significantly correlated with the module interconnectivity matrix ($P < 0.01$) and that they were highly similar (Rand measure = 73%).

Toward Identifying the Signaling Network Involved in Rabies Pathogenesis

In order to retrieve casual relationships, we used the KEGG database and enrichment results to prune the proximal network of the seed gene set. By reviewing the significantly enriched pathways, all KEGG pathways ($N = 47$; **Supplementary Table 3**) were merged to reconstruct the enriched signaling network pertaining to rabies pathogenesis. The full signaling network is presented in **Supplementary Table 5**, but the merge of only 22 of them were presented in **Figure 5**. These 22 pathways were selected based on gene coverage, reported relevance to rabies pathogenesis and association with other viral infections. Then, the DEGs related to these pathways were used to mine the rabies-implicated signaling network (RISN) based on the following KEGG pathways: PI3K-AKT signaling pathway (KEGG:04151), cell cycle (KEGG:04110), Jak-STAT signaling pathway (KEGG:04630), circadian rhythm

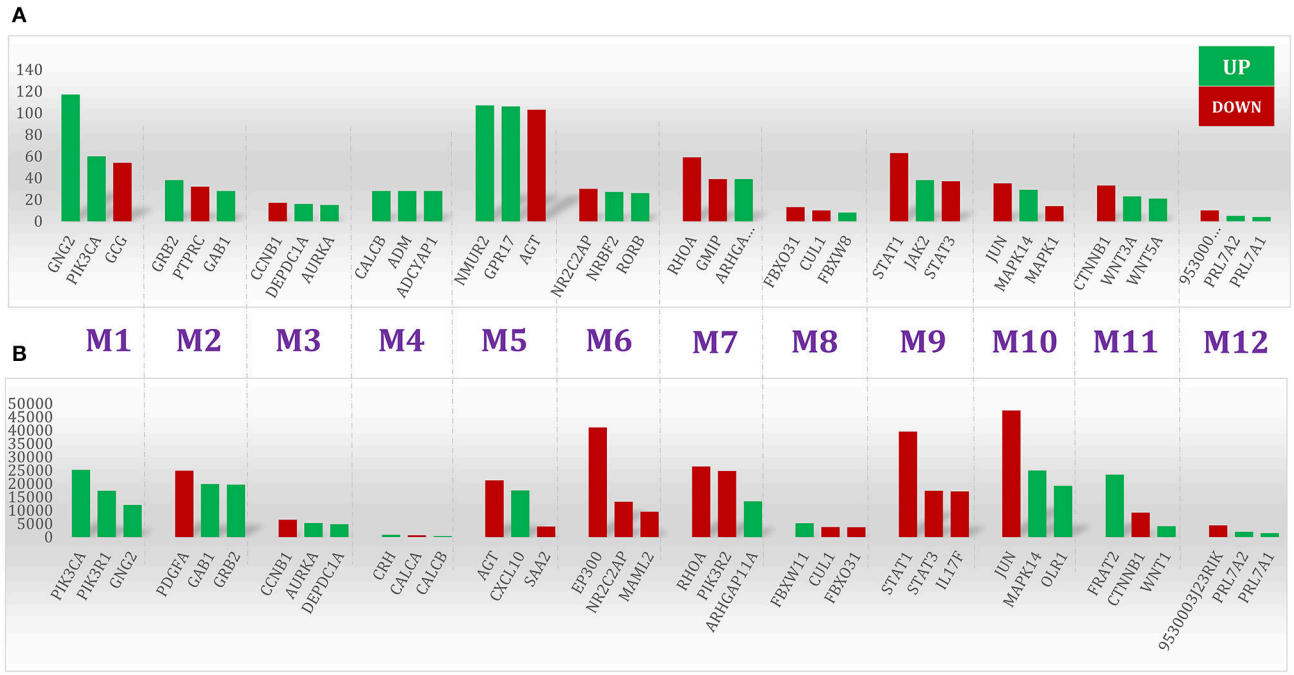


FIGURE 4 | Top ranked nodes based on two centrality measures. The **(A)** degree and **(B)** betweenness centrality measures were calculated in the 12 modules (M1–M12) of the SHIDEG-PPIN proximal neighborhood network of the seed gene set. The over-expression and under-expression of these gene products are labeled by color.

(KEGG:04710), pertussis (KEGG:05133), leishmaniasis (KEGG:05140), tuberculosis (KEGG:05152), hepatitis B (KEGG:05161), influenza A (KEGG:05164), herpes simplex infection (KEGG:05168), Epstein-Barr virus infection (KEGG:05169), inflammatory bowel disease (IBD) (KEGG:05321), PPAR signaling pathway (KEGG:03320), hematopoietic cell lineage (KEGG:04640), neuroactive ligand-receptor interaction (KEGG:04080), Notch signaling pathway (KEGG:04330), inflammatory mediator regulation of TRP channels (KEGG:04750), TNF signaling pathway (KEGG:04668), T cell receptor signaling pathway (KEGG:04660), cytokine-cytokine receptor interaction (KEGG:04060), chemokine signaling pathway (KEGG:04062), and ubiquitin mediated proteolysis (KEGG:04120). The main sink and source nodes in this directed network along with the nodes with high betweenness centrality in the whole RISN are listed in **Table 3**. The influence of nodes with high betweenness on propagating or focusing information among this signaling network is presented by the information release index (IRI), $IRI = \log(Outdegree/Indegree)$. The positive value of IRI indicates the propagating role of nodes and vice versa.

Manually Curated Version of RISN

To simplify RISN, signaling pathways were manually extracted and merged based on the currently available data in KEGG, including WNT, MAPK/ERK, RAS, PI3K/AKT, Toll-like receptor, JAK/STAT, and NOTCH signaling pathways. The information flow from diverse ligands to various transcription factors is illustrated along with differential expression. As shown in **Figure 6**, information is converged toward several important proteins including PLC, MAPK1/2, PIK3, PKC, and JAK, and is then diverged toward several distinct transcription factors and finally end-point biological processes.

Our analysis revealed that two of three WNT signaling pathways were altered in rabies infected cells. The canonical WNT pathway (WNT/β-catenin) along with the non-canonical planar cell polarity (PCP) pathway were apparently active in infected neurons but the non-canonical WNT/calcium pathway was not induced. The PCP pathway is involved in up-regulation of components of the downstream pathway and cytoskeletal rearrangements of which the latter may implicate this pathway in cytoskeletal changes in neurons. This is consistent with previous studies reporting cooperative cytoskeletal changes (restructuration) for viral protein transportation and viral localization (Sagara et al., 1995; Ceccaldi et al., 1997; Song et al., 2013; Zandi et al., 2013).

There is also evidence of crosstalk between WNT and MAPK/ERK signaling pathways. It seems that in the rabid brains the MAPK/ERK signaling pathway, via cAMP-PCREB signaling, is involved in neuromelanin biosynthesis of which its accumulation depletes iron ions as observed in some neurodegenerative diseases such as Parkinson's disease (Good et al., 1992). Iron deficiency may also contribute to defective dopaminergic interaction with neurotransmission systems (Youdim, 2008). This is, however, a speculation and needs experimental validation in rabies infection cases.

Additionally, RAS signaling is activated through the C-Kit receptor and diverge toward PIK3 and MAPK/ERK signaling

TABLE 2 | The enriched KEGG and Reactome pathways.

Module No.	KEGG	Frequency	*P*-value corrected with Bonferroni step down	Reactome	Frequency	*P*-value corrected with Bonferroni step down
M1	Vasopressin-regulated water reabsorption	26	1.13E-33	Peptide ligand-binding receptors	22	9.62E-30
	Calcium signaling pathway	21	2.76E-28	Platelet activation, signaling, and aggregation	10	2.76E-08
	cGMP-PKG signaling pathway	7	2.56E-05	Thrombin signaling through proteinase activated receptors (PARs)	7	2.49E-10
M2	PI3K-Akt signaling pathway	30	2.16E-24	Immune system	49	9.78E-28
	Pathways in cancer	30	1.26E-22	Innate immune system	38	1.31E-25
	Ras signaling pathway	29	2.10E-28	Adaptive immune system	35	6.87E-23
M3	Cell cycle	10	1.71E-11	Cell cycle	28	4.52E-31
	Vasopressin-regulated water reabsorption	2	2.42E-02	Cell cycle, mitotic	27	4.64E-31
	Bladder cancer	2	4.17E-02	M Phase	12	1.80E-10
M4	Neuroactive ligand–receptor interaction	16	3.08E-22	G alpha (s) signaling events	23	1.04E-47
	–	–	–	GPCR ligand binding	21	1.91E-29
	–	–	–	Class B/2 (secretin family receptors)	10	5.73E-16
M5	Chemokine signaling pathway	25	1.78E-29	Signaling by GPCR	68	2.40E-78
	Neuroactive ligand-receptor interaction	25	2.85E-25	GPCR downstream signaling	63	2.90E-70
	Cytokine-cytokine receptor interaction	20	4.75E-18	G alpha (i) signaling events	63	1.24E-110
M6	Thyroid hormone signaling pathway	12	1.56E-12	Generic transcription pathway	46	3.10E-54
	Notch signaling pathway	8	8.68E-10	Developmental biology	28	1.47E-24
	Maturity onset diabetes of the young	3	7.71E-03	Nuclear receptor transcription pathway	27	4.31E-53
M7	Regulation of actin cytoskeleton	9	1.74E-06	Signaling by Rho GTPases	52	1.49E-77
	Pancreatic cancer	3	3.45E-02	Rho GTPase cycle	51	1.99E-104
	–	–	–	G alpha (12/13) signaling events	17	1.45E-25
M8	Ubiquitin mediated proteolysis	8	2.36E-12	Association of TriC/CCT with target proteins during biosynthesis	4	1.40E-07
	Circadian rhythm	2	1.19E-03	Protein folding	4	1.44E-06
	–	–	–	Chaperonin-mediated protein folding	4	1.04E-06
M9	Jak-STAT signaling pathway	37	2.06E-50	Immune system	57	3.06E-41
	Cytokine–cytokine receptor interaction	36	1.80E-39	Cytokine signaling in immune system	53	1.54E-67
	Measles	26	9.44E-32	Interferon signaling	31	1.08E-35
M10	MAPK signaling pathway	25	3.59E-24	Innate immune system	31	2.54E-21
	Pathways in cancer	23	7.36E-17	Toll-like receptors cascades	19	1.54E-20
	PI3K-Akt signaling pathway	22	7.01E-17	Toll like receptor 3 (TLR3) cascade	18	7.97E-22
M11	Wnt signaling pathway	33	6.80E-57	Signaling by Wnt	30	1.06E-37
	Pathways in cancer	28	2.05E-31	TCF dependent signaling in response to WNT	23	1.72E-28
	Melanogenesis	26	2.21E-44	Class B/2 (secretin family receptors)	18	1.63E-27
M12	–	–	–	Amyloids	5	1.21E-09
	–	–	–	Disease	5	8.05E-05

The statistically significantly enriched KEGG and Reactome pathways were identified by ClueGO. The top three representative pathways identified in each module (M1–M12) of the SHIDEG-PPIN proximal neighborhood network are given together with their corrected p-values. The highlighted pathway names were found to be enriched in more than one module.

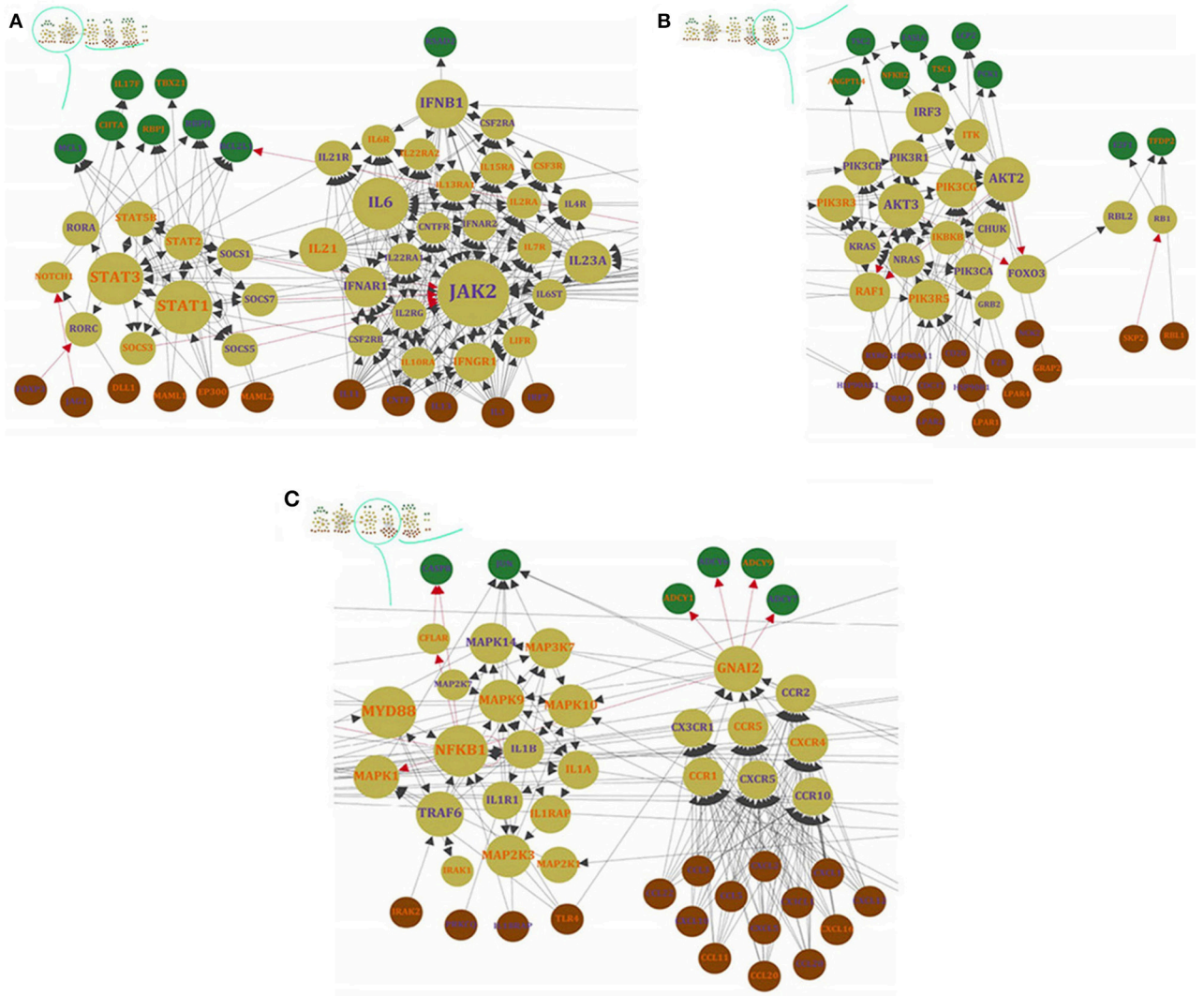


FIGURE 5 | A bird's eye view of RISN. The 22 over-expressed genes of significantly enriched KEGG pathways were selected and merged. The DEGs are labeled by different colors including orange and blue which indicates down-regulation and up-regulation of genes, respectively. The activator/inhibitor edges are also colored differently (red edges are inhibitors and black vice versa). The node colors represent its relative position in a directed network from brown (source nodes), cream (internal nodes) to green (sink nodes) and the node size is proportional to betweenness centrality value. In addition, the six detected modules in three parts are displayed separately: **(A)** Interferon circumvent, **(B)** Toward proliferation and survival, and **(C)** neuropathological clue.

pathways. Downstream of RAS activation (ERK signaling and AKT) is highly complex but generally contributes to cell growth (Bender et al., 2015). Activated RAS signaling suppresses PKR-mediated responses to interferon response and double-stranded RNA degradation. Normally, viral transcripts trigger PKR phosphorylation and activation, and finally inhibit infection. Therefore, the RABV may replicate silently in RAS activated cells (Mundschau and Faller, 1994; Russell, 2002). This data-based hypothesis also requires experimental validation in rabid cases.

The AKT signaling pathway plays a critical role in the replication of the RABV similar to other non-segmented negative-stranded RNA viruses. Heavy phosphorylation of viral proteins (P protein) is mainly mediated via AKT activity (Sun et al., 2008). Subsequently, the activated P protein plays a crucial role in other signaling pathways such as Toll-like receptor and JAK/STAT signaling pathways which are responsible for viral genome detection and immune-modulatory functions against rabies, respectively. Accordingly, the viral G protein activates AKT signaling through phosphorylation and localization of PTEN (Terrien et al., 2012). The consequences of the activation and crosstalk of these signaling pathways are reduced apoptosis, cell survival and blocked cell cycle progression. Neuronal dysfunction, inhibition of apoptosis, and limitation of inflammation have been previously stressed by Gomme et al. (2012). It seems that these processes have been evolutionarily acquired to complete virus lifecycle and transfer to the new host. They also showed that most of DEGs are involved in signaling transduction and nervous system function, and therefore affect cell behavior by decreasing neurite growth,

TABLE 3 | Details of the main sink and source nodes along with high betweenness centrality values in the whole RISN.

Gene symbol	Protein name	Po.	U/D	Role in rabies
CCL11	Chemokine (C-C motif) ligand 11	Source	Down	
CCL20	Chemokine (C-C motif) ligand 20	Source	Down	
CCL22	Chemokine (C-C motif) ligand 22	Source	Up	
CCL28	Chemokine (C-C motif) ligand 28	Source	Up	
CCL3	Chemokine (C-C motif) ligand 3	Source	Up	Highly activated in the brains of mice infective with Rabies (71)
CCL5	Chemokine (C-C motif) ligand 5	Source	Up	Known as a vital regulator which is involved in convincing encephalomyelitis (72), raised levels of mRNA transcripts (73), the expression value of CXCL10 and CCL5 in microglia is accurately regulated while the multiple signaling pathways are activated
CD28	CD28 molecule	Source	Up	
CDC37	Cell division cycle 37	Source	Up	
CNTF	Ciliary neurotrophic factor	Source	Up	
CX3CL1	Chemokine (C-X3-C motif) ligand 1	Source	Up	
CXCL1	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	Source	Up	
CXCL10	Chemokine (C-X-C motif) ligand 10	Source	Up	Known as a vital regulator which is involved in convincing encephalomyelitis (72), raised levels of mRNA transcripts (73), the expression value of CXCL10 and CCL5 in microglia is accurately regulated while the multiple signaling pathways are activated
CXCL12	Chemokine (C-X-C motif) ligand 12	Source	Up	
CXCL16	Chemokine (C-X-C motif) ligand 16	Source	Down	
CXCL2	Chemokine (C-X-C motif) ligand 2	Source	Up	
CXCL5	Chemokine (C-X-C motif) ligand 5	Source	Up	
DLL1	Delta-like 1 (Drosophila)	Source	Down	
HSP90AA1	Heat shock protein 90 kDa alpha family class A member 1	Source	Up	Associated with virus packaging (68)
HSP90AB1	Heat shock protein 90 kDa alpha family class B member 1	Source	Up	Associated with virus packaging (68)
HSP90B1	Heat shock protein 90 kDa beta family member 1	Source	Up	Associated with virus packaging (68)
IL11	Interleukin 11	Source	Up	
IL3	Interleukin 3	Source	Up	
JAG1	Jagged 1	Source	Up	
LPAR1	Lysophosphatidic acid receptor 1	Source	Down	
LPAR2	Lysophosphatidic acid receptor 2	Source	Up	
LPAR4	Lysophosphatidic acid receptor 4	Source	Down	
MAML1	Mastermind like transcriptional coactivator 1	Source	Down	
MAML2	Mastermind like transcriptional coactivator 2	Source	Down	
ANGPTL4	Angiopoietin like 4	Sink	Down	
BCL2L1	BCL2-like 1	Sink	Down	
IL17F	Interleukin 17F	Sink	Down	
MCL1	Myeloid cell leukemia 1	Sink	Up	
PCK1	Phosphoenolpyruvate carboxykinase 1 (soluble)	Sink	Up	
IL21	Interleukin 21	1.62	Down	IL-21 is critical for the development of optimal vaccine-induced primary but not secondary antibody responses against RABV infections (Dorfmeier et al., 2013)
IFNB1	Interferon, beta 1, fibroblast	1.19	Up	RABV P protein binds and inhibit Binding to IRF3 (Brzózka et al., 2006)
EP300	E1A binding protein p300	0.70	Down	
TLR4	Toll-like receptor 4	0.64	Down	No sign of phenotype due to lacking TLR4
IL6	Interleukin 6	0.58	Up	Overexpression during infection (58,59), correlation among the IL-6 genes and the way of behavioral lateralization (60), involved in RABV pathogenesis (61)

(Continued)

TABLE 3 | Continued

Gene symbol	Protein name	Po.	U/D	Role in rabies
MYD88	Myeloid differentiation primary response 88	0.44	Down	Weakened RABV intervenes deadly disease while no MyD88 present, genetic adjuvanting with Myd88 improved the RVNA responses of a plasmid DNA rabies vaccine (90,91)
NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	0.13	Down	
STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)	0.12	Down	inhibits STAT3 nuclear accumulation (Lieu et al., 2013)
TRAF6	TNF receptor associated factor 6	0.08	Up	
F2R	Coagulation factor II (thrombin) receptor	0.05	Up	
CASP8	Caspase 8, apoptosis-related cysteine peptidase	0.00	Up	Activation in RABV (Sarmento et al., 2006)
GNAI2	Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2	0.00	Down	
RBL1	Retinoblastoma-like 1	0.00	Down	
TFDP2	Transcription factor Dp-2 (E2F dimerization partner 2)	0.00	Down	
STAT1	Signal transducer and activator of transcription 1	−0.01	Down	RABV P protein binds and inhibit dimerization of STAT (Vidy et al., 2005; Brzózka et al., 2006), inhibit translocation (Brzózka et al., 2006; Moseley et al., 2009)
AKT2	v-AKT murine thymoma viral oncogene homolog 2	−0.07	Up	Hyper-phosphorylation of RABV P protein (Sun et al., 2008)
AKT3	v-AKT murine thymoma viral oncogene homolog 3	−0.07	Up	Hyper-phosphorylation of RABV P protein (Sun et al., 2008)
JUN	Jun proto-oncogene	−0.32	Up	Activated in Rabies (Nakamichi et al., 2005)
FASLG	Fas ligand (TNF superfamily, member 6)	−0.48	Up	Immune disruptive strategy of RABV to bring about apoptosis in T cell by overexpression in neuron (93)
IRF7	Interferon regulatory factor 7	−0.60	Up	RABV P protein averts tis activation (70)
ADCY6	Adenylate cyclase 6	−0.70	Up	The signal pathway from the stimulating regulatory component of the adenylate cyclase system to the unchanged activity of the catalytic subunit is defective (Koschel and Halbach, 1979; Koschel and Münzel, 1984)
ADCY1	Adenylate cyclase 1 (brain)	−0.73	Down	The signal pathway from the stimulating regulatory component of the adenylate cyclase system to the unchanged activity of the catalytic subunit is defective (Koschel and Halbach, 1979; Koschel and Münzel, 1984)
ADCY7	Adenylate cyclase 7	−0.73	Up	The signal pathway from the stimulating regulatory component of the adenylate cyclase system to the unchanged activity of the catalytic subunit is defective (Koschel and Halbach, 1979; Koschel and Münzel, 1984)
ADCY9	Adenylate cyclase 9	−0.73	Down	The signal pathway from the stimulating regulatory component of the adenylate cyclase system to the unchanged activity of the catalytic subunit is defective (Koschel and Halbach, 1979; Koschel and Münzel, 1984)
IRF3	Interferon regulatory factor 3	−0.78	Up	RABV P protein binds and avert binding to IFNB1 (53), inhibit IRF3 phosphorylation (70,96)
JAK2	Janus kinase 2	−0.79	Up	

The third and fourth column indicate the position (Po.) of the corresponding genes in RISN and expression changes (U/D) of them based on our meta-analysis. The IRI of the nodes with high betweenness values are presented in the third column

organization of cytoskeleton and cytoplasm, and microtubule dynamics.

It has been demonstrated that rabies infection up-regulates expression of CXCL10 and CCL5 proteins in a ERK1/2-, p38-, and NFkB-dependent manner (Nakamichi et al., 2004, 2005). CXCL10 is a major chemo-attractant of Th-1 cells. The up-regulation of interferon, chemokines, interleukin (IL), and IL-related genes were previously observed by Sugiura et al. (2011). They also reported the signaling pathways involved in rabies infection including interferon signaling, IL-15 production and signaling, and Granzyme B signaling which trigger apoptosis in immune target cells.

ERK and p38MAPK along with BCL2 family and the FasL receptor are important apoptosis committers. Currently, it is known that RISN inhibits apoptosis and also suppresses cell proliferation (Gomme et al., 2012). Also, JAK/STAT signaling is activated through its receptors, but as mentioned earlier STAT dimerization is inhibited via viral P protein activity. Therefore, downstream signaling of STATs, which is critical for interferon signaling and viral defense, is suppressed. Concomitant activation of JAK/STAT and AKT signaling

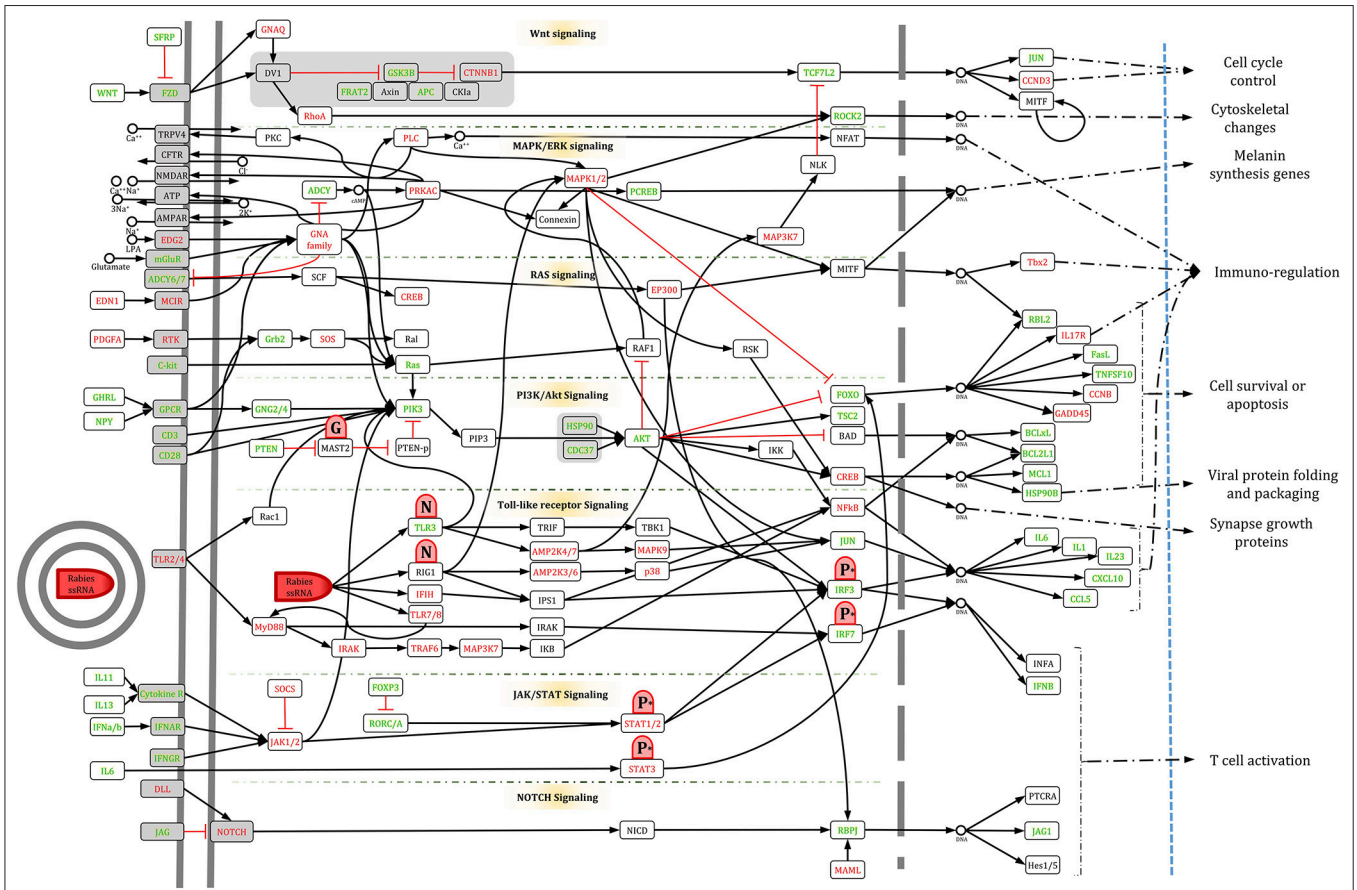


FIGURE 6 | The manually curated version of RISN. The over- and under-expressed genes are colored green and yellow, respectively. The cellular membrane and nuclear membrane are depicted by gray solid and dashed lines, respectively. Proteins are depicted by rounded rectangles, and DNA and small metabolite molecules are shown by circle. The phosphorylated P protein activated by AKT is denoted by *P**, however, for the sake of simplicity, its process is not shown in the signaling pathway. Viral components are depicted by red shapes but with different caps to delineate activatory and inhibitory effects of RABV on RISN.

pathways has a pro-survival function in neurons (Junyent et al., 2010). In a time-course study, Zhao et al. studied the gene expression profile of infected microglial cells and indicated some affected signaling pathways at different time points (Zhao et al., 2013). The MAPK, chemokine, and JAK-STAT signaling were also shown to be implicated in rabies infection along with other innate and adaptive immune response pathways. These pathways were also detected in two other independent studies on CNS of infected mice (Zhao et al., 2011, 2012b).

Viral pattern recognition is critical for early innate immunity response and modulation of pathogen-specific adaptive immunity. TLR3, a member of the TLR family which are pattern recognition receptors in cells, is increased in the cytoplasm of rabies infected cells. It plays major functions in spatial arrangements of infected cells and viral replication, and is observed in endosomes and Negri bodies which are only formed in the presence of TLR3 (Ménager et al., 2009).

NOTCH signaling is important for cell communication, neuronal function, and development in spatial learning and memory (Costa et al., 2003). Our data indicate that this signaling pathway is active in rabid brains. More detailed examination of the role of this signaling pathway in rabies infection is warranted.

Experimental Validation of Expression Alterations

Gene expression analysis of a number of randomly selected DEGs has been performed previously based on RT-qPCR as a routine validation method of microarray-based expression profiles (Zhao et al., 2011, 2012a,b, 2013) or fluorescent bead immunoassay (Sugiura et al., 2011). In all cases, the former experimentally confirmed DEGs identified in SHIDEGs were *STAT1*, *STAT3*, *SOCS2*, *IRF1*, *IRF3*, *IRF7*, *IFNAR2*, *SH2D1A*, *CCL3*, *CCL5*, *CCRL2*, *CXCL10*, *Mx1*, *IFIT3*, *OASL2*, *USP18*, *IL6*, *IL10*, *IL23A*, and *RTP4*. However, we examined another independent random gene set among RISN genes as a further step of validating the microarray-based results. The differential expression of *AKT3*, *GNAI2*, and *IL21* was analyzed by comparing expression levels in murine neuroblastoma cells infected by the wild type RABV with control uninfected cells using RT-qPCR. The results indicated that the direction of differential expression

of all three genes were consistent between RT-qPCR results and data integrated from multiple microarray chips (**Figure 7**, **Supplementary Tables 6**, **7**). These results confirmed the down-regulation of *GNAI2* and *IL21*. *AKT3* expression values and statistical tests state that the expression of *AKT3* is not up-regulated in infected samples. Our findings based on delta *Ct* method and comparing the raw expression values of *AKT3* gene in both samples with the referenced values; however, firmly confirm that the value of gene is indeed up-regulated in infected samples.

DISCUSSION

Rabies is a fatal neuropathological disorder. The fatality of this infection is not because of neurological damage or neurohistopathological signs, but due to neurophysiological disruption of vital signs such as regular heart beat and respiratory rhythm. Other evidences that highlight this neural malfunction are known rabies symptoms such as hydrophobia, photophobia, and paralysis of facial and throat muscles. Although rabies infected cells can mount an innate immune response against this infection, the virus can control the expression and function of the proteins involved in the induction of apoptosis and efficiently suppresses the antiviral innate immune response. From a pathobiological point of view, we acknowledge that the RISN and the previously reported pathways which lead to the spread of RABV can also be triggered by unrelated viruses including other neurotropic RNA viruses, measles, and influenza. This infection would yield similar gene expression profiles by activating general host responses including activation of stress response, innate immune response, and interferon signaling signatures. Based on biological relevance, we partitioned RISN under the following three functional domains.

Interferon Circumvent

To escape innate and adaptive immunity, rabies perturbs Jak-Stat signaling by influencing interactions and expression. As shown in **Figure 5A**, two modules of RISN are likely to interfere with this signaling pathway. The inhibition of dimerization of Stat proteins and accumulation in nucleus by viral P protein are previously described (Vidy et al., 2005; Brzózka et al., 2006; Moseley et al., 2009; Lieu et al., 2013). Our findings showed the

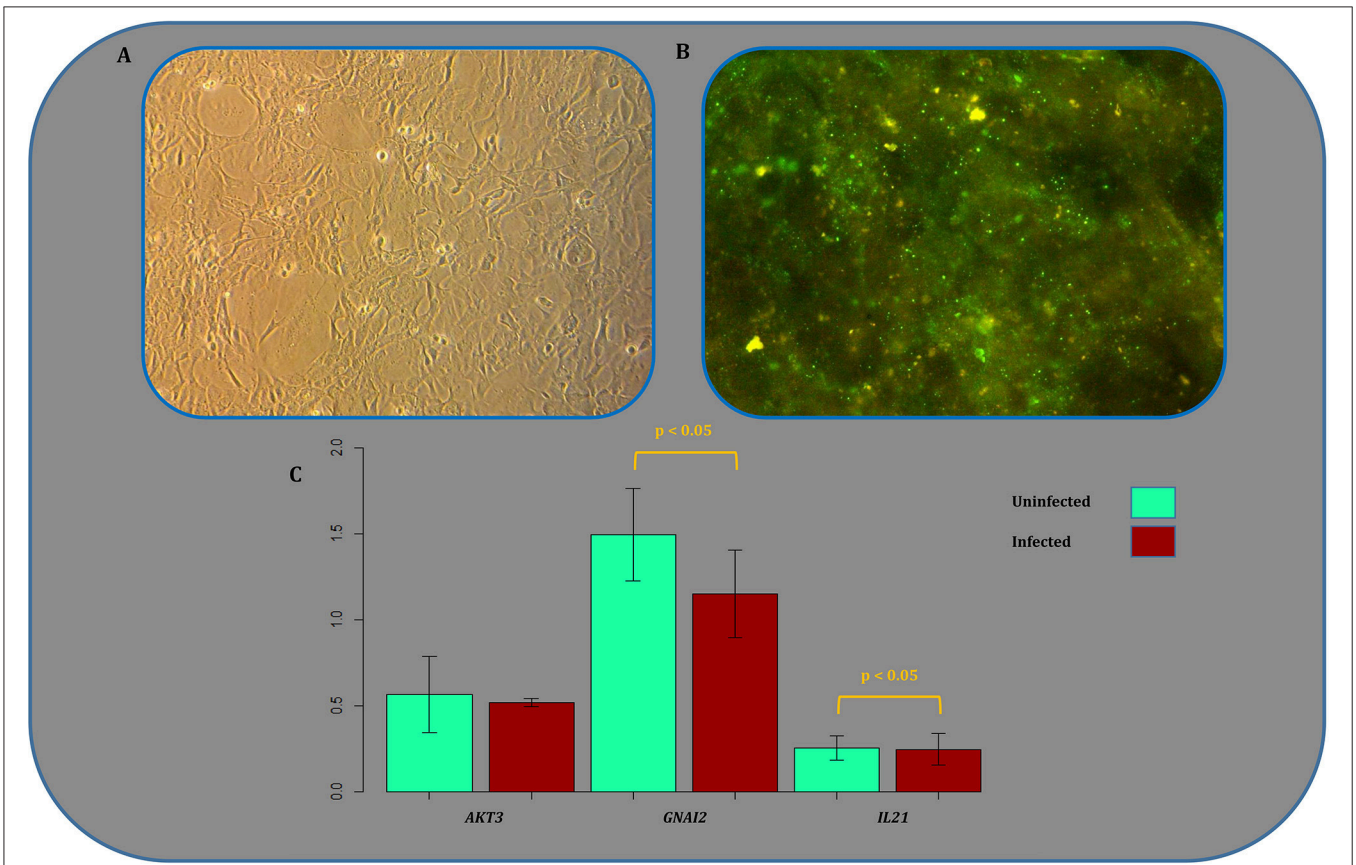


FIGURE 7 | Experimental validation of microarray-based expression results. (A) Mock infected N2a cell culture image showing a normal morphology. **(B)** N2a cells infected with CVS strain of RABV with multiplicity of infection (MOI = 3), stained by FITC conjugated anti-rabies nucleocapsid polycolonal antibody. The images were captured 24 h post infection. **(C)** Expression profile of three randomly selected genes acting in different signaling pathways of the rabies-implicated signaling network (RISN). Expression levels were quantified using RT-qPCR in triplicates. The error bar plots indicate Mean ± *SD* and include the corresponding *p*-value of the statistical significance test. The vertical axis measures the negative inverse value of the logarithm of the mean value for each replicate using the delta *Ct* method.

down-regulation of STAT proteins, which is plausible considering the feedback self-loop control on these proteins. This finding is supported by up-regulation of the feedback inhibitor known as the SOCS protein. The high betweenness value of JAK2 and its up-regulation indicate the role of activating innate and adaptive defense systems of infected cells against rabies. Additionally, the low IRI-value of JAK2 and transfer of information toward the "toward proliferation and survival" modules highlight its importance in rabies pathogenesis.

The expression alterations of IL6 and IL21 have been reported in rabies infection previously (Hemachudha et al., 1993; Megid et al., 2006; Quaranta et al., 2008; Dorfmeier et al., 2013; Srithayakumar et al., 2014). The down-regulation of IL21 and its receptors, IL21R, IL15R, and IL17F, following the down-regulation of STAT3, is very significant in the observed dampened immune system of the rabid given their role in proliferation and maturation of natural killer cells. Unlikely, the up-regulation of IL6, a neuroprotective cytokine, reinforces the anti-apoptotic effects of the rabies wild type strain. Both of these play a propagation role in this network with IRI-values above one. Despite vastly interfering with the Jak-Stat signaling, rabies infection could not decrease the expression of the famous antiviral molecule, IFNB1, and cell could have upregulated it against the infection. Surprisingly, however, the virus chooses another strategy to skip the interferon mechanism (Faul et al., 2010). This alternative plan is to perturb IFNB1 activation via IRF3. The over-expression of RORA and RORC is also important since they affect the circadian rhythm, calcium-mediated signal transduction and anti-inflammatory responses. It seems that the up-regulated MCL1 and BCL2L1, act in favor of survival, inflammation attenuation and apoptosis inhibition of neurons and disrupt endocytic vesicle retrieval. Also, decrease in function of CIITA and TBX1 causes decrease in the function and efficiency of TH1 and TH2. Overall under-expression of the Notch signaling pathway, including DLL1, NOTCH1, and RBPJ, along with the up-regulation of JAG1, an inhibitor of NOTCH1, is indicative of malfunction in cell-cell communication in CNS and neuronal self-renewal mechanism.

Toward Proliferation and Survival

Similar to other viral infection, the RABV hijacks the proliferation machine of cells to generate virions as much as possible. To achieve this goal, the strategy of the virus is to keep the cell alive and active and amplifies the production rate. Rabies achieves this by using cell envelope and preventing apoptosis. In fact, there was an inverse correlation between induction of apoptosis and the potency of a virus strain to invade the brain. This suggests that suppression of apoptosis may well be a strategy for neuro-invasiveness of pathogenic RABV and progression through the nervous system (Thoulouze et al., 2003a,b; Larrous et al., 2010). As shown in **Figure 5B**, up-regulated AKT2 and AKT3 play a central role in the tyrosine kinase module. It has been previously reported that AKT signaling is hijacked by non-segmented RNA viruses such as vesicular stomatitis virus (VSV) via phosphorylating P proteins (Sun et al., 2008). It has also been suggested to use AKT inhibitors as an anti-RABVdrug. Our findings underscore the importance of this signaling pathway in neuronal cells where AKT signaling is not normally hyperactive. This activated pathway, prompted by G viral proteins, may result in the activation of proliferation and growth machinery, and help viral protein folding and packaging via over-expression of HSPs (Lahaye et al., 2009, 2012). TSC2 also triggers apoptosis in immune cells via high representation of FASLG. These results in parallel with those in Sun et al. (2008) highlights the need for studying AKTs and anti-AKTs in rabies models.

Other Serine or Tyrosine kinases, including ITK, IKBKB and RAF1, and PIK3CG were under-expressed, thus resulting in the dampening of the inflammatory response especially with the up-regulation of anti-inflammatory proteins such as CHUK and cell proliferatory proteins, NRAS and KRAS. IRF3, IRF7, and IFNB1, which are upregulated naturally in response to viral infection, could not stimulate an immune response due to the activation of the P viral protein. Besides, IRF3 and IRF7 are involved in AKT activation and transformation of inflammatory to anti-inflammatory macrophages (Rieder et al., 2011; Tarassishin et al., 2011). The over-expressed AKT genes also inactivate FOXO3 and therefore disrupt the cell efforts toward apoptosis (Tarassishin et al., 2011). Expression of ANGPTL4 that causes Anoikis, a type of programmed cell death, is also decreased after this module activity (Terada and Nwariaku, 2011). The expression alteration of proteins involved in cell cycle regulation including RB1, RBL1, RBL2, TFDP2, and E2F1 is indicative of the triggering disruption and hijacking by the virus.

Neuropathological Clue

Hitherto, the underlying mechanism of escape from the immune system, apoptosis prevention and virus production in rabies infection was demonstrated by these modules (**Figures 5A,B**). However, importantly, the main cause of death in rabid is cardiac arrhythmia and breathing pattern disorders, for which its molecular basis should be tracked elsewhere. Meanwhile, trace of chemokines such as CCL3, CCL5 (a neuron survival factor), and CXCL10 in rabies infection has been previously detected (Nakamichi et al., 2005; Johnson et al., 2008; Li et al., 2012; Huang et al., 2014). These molecules along with their receptors are involved in the blood brain barrier (BBB) permeability and recruitment of different T cells (**Figure 5C**). In the natural cell cycle, MYD88 expression leads to an increase in expression of NFκB and hence programmed cell death. Seemingly, the expression and function of MYD88 in addition to NFκB is decreased (**Figure 5C**) which can lead to the suppression of apoptosis. Diverse chemokines as source nodes in RISN and information transduction to adenylate cyclases and MAP kinases via GNAI2 is likely to be a critical clue to discovering the etiological mechanism of rabies fatality.

Adenylyl cyclases (ADCYs) are central components of signaling cascades downstream of many G proteins. In the mammalians, of the ten ADCY isoforms identified, nine (ADCY1-9) are transmembrane proteins, whereas ADCY10 is a soluble isoform that lacks the transmembrane domains (Sunahara et al., 1996; Conley et al., 2013; Birrell et al., 2015). Although in the RISN the expression of *ADCY1* and *ADCY9* was decreased, the level of ADCY6 and ADCY7 was increased, the overall outcome is probably reduction of signal flow rousted by

GPCR and ADCY. All ADCY isoforms catalyze the conversion of ATP to cyclic AMP (cAMP) and pyrophosphate. cAMP is a messenger involved in many biological processes including cell growth and differentiation, transcriptional regulation, apoptosis, and various other cellular functions (Patel et al., 2001). The main protein kinase activated by cAMP is protein kinase A (PKA). PKA transfers phosphate groups form ATP to proteins including ion channels on the cell membrane. Similar to changes in enzyme activity following biochemical modification, phosphorylation of ion channel proteins may also cause conformational changes and consequently increase chances of channel opening leading to depolarization of postsynaptic neurons, resulting in firing an action potential and altered electrical activity properties. On the other hand, ADCYs are integrated in lipid rafts and caveolae, and implicated in local cAMP micro-domains in the membrane (Schwencke et al., 1999). Subcellular compartmentalization of protein kinases (such as PKA) and phosphatases, through their interaction with A kinase anchoring protein (AKAPs), provides a mechanism to control signal transduction events at specific sites within the cell (McConnachie et al., 2006). The RABV may interfere with the lipid raft and the micro-compartment associated with the cAMP–AKAP–PKA complex and thus alter ion channel function, eventually leading to neuronal dysfunction. In line with this, it has been reported that NMDA and AMPA glutamate receptors form complexes with cytoskeletal and scaffold proteins in the post-synaptic density (PSD; Kennedy, 1997; Ziff, 1997). Interestingly, AKAP binds to PSD in complexes with NMDA and AMPA receptors (Colledge et al., 2000). It is also thought that regulation of this molecular architecture is essential for controlling glutamate receptors in hippocampal long-term potentiation (LTP) and long-term depression (LTD) synaptic plasticity (Lüscher et al., 2000; Tomita et al., 2001). Our data showed that the expression of PRKACA and AKAP13 subsequent to GNAI2 decreased significantly (**Supplementary Table 5**). It therefore seems that these complexes may be a preferential target of viruses to hijack cellular machinery.

Neuropathological observations indicate that functional alterations precede neuronal death, which is responsible for the clinical manifestation and fatal outcome in rabies. Indeed, Gourmelon et al. reported that disappearance of rapid eye movement (REM) sleep and the development of pseudoperiodic facial myoclonus are the first manifestations in the EEG recordings of mice infected with the challenge virus standard (CVS) of fixed RV (Gourmelon et al., 1986). It has also been reported that electrical activity of brain terminates 30 min prior to the cardiac arrest, indicating that cerebral death occurs before vegetative function failure in experimental rabies (Fu and Jackson, 2005). Considering the increased activity of voltage-gated channels by phosphorylation in response to PKA stimulation, initiating a signaling pathway from ADCY to ion channel functioning could be a possible mechanism by which the RABV hijacks the neurons. Consistently, Iwata et al. showed that ion channel dysfunction occurs in mouse neuroblastoma cells infected by RV (Iwata et al., 1999). They reported that not only the functional activity of voltage-dependent sodium and inward rectifier potassium channels were decreased, but also the resting membrane potential was decreased, indicating membrane depolarization. Therefore, decreased activity of these channels could preclude infected neurons from firing action potentials and generating synaptic potentials, thus leading to functional impairment. Fu and Jackson (2005) observed that neurotransmitter releases from rat hippocampus, after inoculation with RV CVS-24, was increased at day 1, reached a peak at day 3, and then declined by day 5. Manifestations of clinical signs of rabies were consistent with day 5 of inoculation when neurotransmitter release was equal or below the level prior to infection, suggesting that neurons are no longer capable of releasing neurotransmitters at the synaptic junctions and this may be the underlying basis of clinical signs including paralysis (Fu and Jackson, 2005).

Since there is a paucity of data pertaining to rabies influence on physiological processes, particularly on neuronal electrophysiological properties, further studies need to be undertaken to confirm whether neuronal dysfunction occurring in rabies infection is due to an aberrant signaling pathway initiating from ADCY-cAMP-PKA and finalizing with ion channel phosphorylation. It should be noted that any alterations in different ion channels may result in dysfunction of neurons and brain regions which are responsible for vital tasks including attention, thinking and respiration.

CONCLUSION

Knowledge-driven studies are mostly non-automatic, heuristic, expert-dependent, and evidence-based surveys. Although this strategy of problem solving is valuable in identification of novel findings, it suffers from certain limitations and subjective biases. With the emergence of omics technologies, data-driven studies have exploited the large and ever-growing publicly available deposited datasets as a complementary approach to knowledge-based studies (Sun et al., 2012). Data-driven approaches are computationally demanding and require complex interpretations but this is dependent directly to the original data itself (Hua et al., 2006; Sun et al., 2012). Here, we combined data- and knowledge-driven studies to potentially identify a less-biased signaling network of rabies infection. We thus undertook a systematic approach which initiated with a data-driven approach and was then extended by a comprehensive complementary knowledge-based approach. In addition, we included all available high-throughput whole-transcriptome datasets in a horizontal (meta-analysis) and super-horizontal (miRNA and mRNA) integration approach. Critically, signaling pathways were then used as a scaffold for data integration to identify key players in signaling pathways and genes. Finally, we constructed a bird's-eye view map, RISN, of signaling deviations including host–pathogen interaction data. Uniquely, this signaling pathway illustrates the host-rabies interaction signature.

In summary, we demonstrate that seven signaling pathways including (1) WNT, (2) MAPK/ERK, (3) RAS, (4) PI3K/AKT, (5) Toll-like receptor, (6) JAK/STAT, and (7) NOTCH are involved in controlling cell cycle, cell survival, viral replication and folding, synapse regulation, and regulation of immunity. Among the many involved proteins, divergence and convergence

of signals indicates that PLC, MAPK1/2, PIK3, PKC, and JAK are potentially the most critical of all in rabies pathogenesis. Interestingly, signals are converged toward these proteins and are then diverged toward several distinct transcription factors and end-point biological processes (**Figure 8**).

In addition to confirming former reports on the inhibition of apoptosis in neurons, RISN provided molecular evidence of interferon escape and neural cell death prevention in rabies infection. This finding is significant given that it explicates how the virus continues to parasitically multiply without any neural host cell damage. Data herein suggest that, the RABV hijacks the phosphorylation machinery of the cell to facilitate its own replication. Also, the tight regulation of recruited immune cells by the virus is demonstrated. The network analysis also

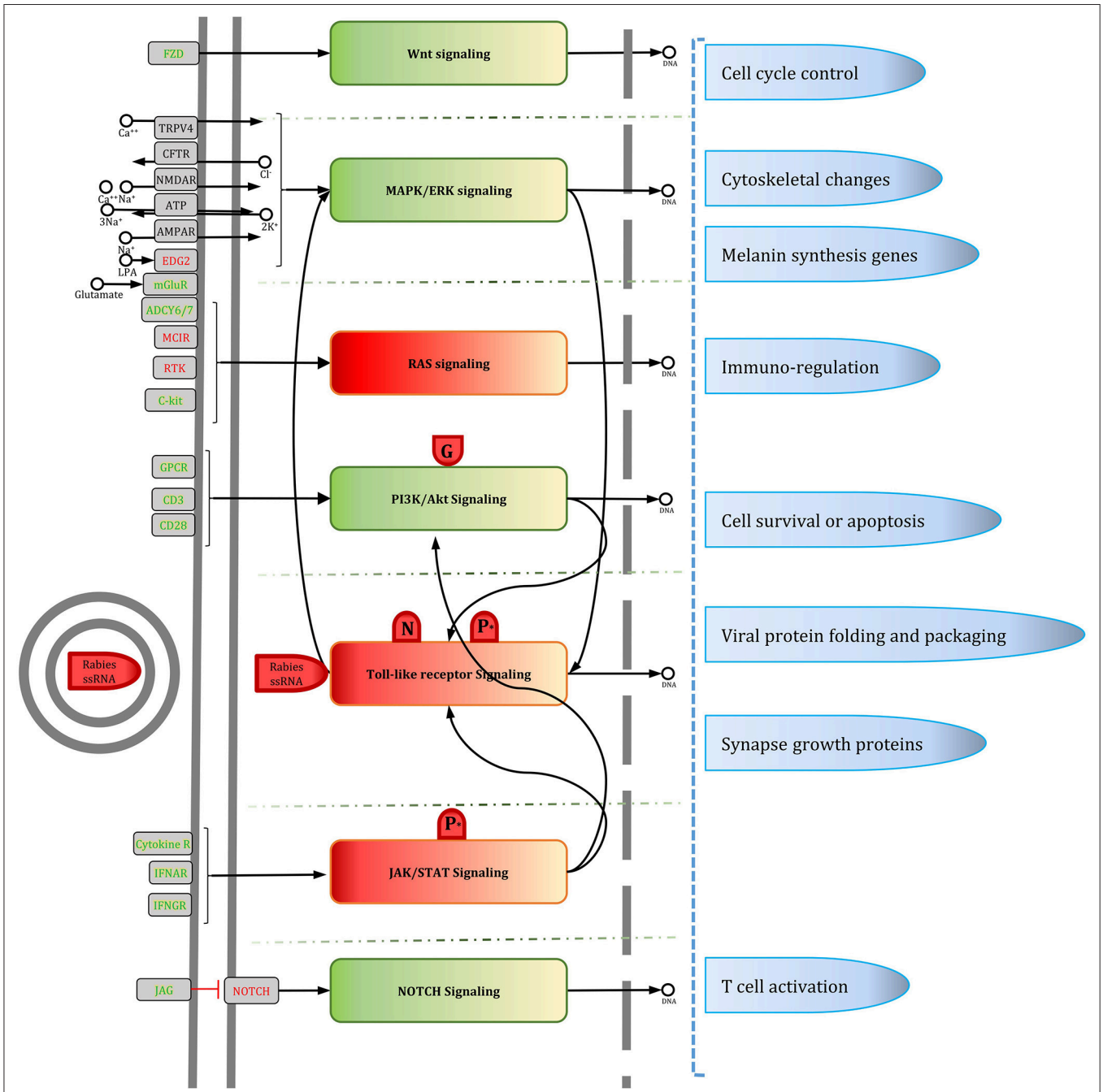


FIGURE 8 | The manually curated version of RISN in a nutshell. To summarize the biological processes involved in the infected neuron, the signaling pathways along with their triggers and consequences are delineated based on **Figure 6**. Activated and inhibited pathways are colored green and red, respectively. Viral components are depicted by red shapes but with different caps to delineate activatory and inhibitory effects of RABV on RISN.

shed light on the gene set central to rabies infection, all of which were bottlenecks in RISN. Moreover, based on RISN, we hypothesize that modifying certain signal transduction apparatus involved in rabies pathogenesis such as the cAMP or AKT signaling pathway may instigate an effective immune response which will consequently diminish the fatality of the rabies infection.

The systems biomedicine approach employed in this study provided a better understanding of the underlying signaling network of this infectious disease. Further independent validation of the RISN potentially provides a molecular framework for intervention and development of novel effective treatments for the late stages of this neglected disease.

AUTHOR CONTRIBUTIONS

MJ conducted the design of study and carried out literature review, data collection, data analysis, and implemented the computational methods. SHM did the literature search and some analysis. MJ, SAJ, and HP participated in network analysis. AG and NA performed the experimental validation. MJ, NA, AG, SAJ and SHM wrote the paper. MM, FN, and BV participated in revising the manuscript critically. All authors read and approved the final manuscript.

FUNDING

This work was supported with a research grant received from Pasteur Institute of Iran (No. 748).

ACKNOWLEDGMENTS

The authors wish to thank Karl-Klaus Conzelmann, Monique Lafon, and Hervé Bourhy for providing us with in-depth knowledge on rabies and their critical evaluation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.01688/full#supplementary-material

Supplementary Table 1 | The selected datasets for rabies data integration.

Supplementary Table 2 | The gene ontology enrichment analysis of differentially expressed and central (degree and betweenness) genes in SHIDEG-PPIN.

Supplementary Table 3 | The full results of pathway enrichment analysis of SHIDEG-PPIN modules by ClueGO.

Supplementary Table 4 | The KEGG enriched pathway similarity matrix and the module interconnectivity matrix of RISN.

Supplementary Table 5 | Properties of RISN nodes and list of all edges.

Supplementary Table 6 | The over- and under-expression of 694 DEGs in the unrefined RISN.

Supplementary Table 7 | Properties of refined SHIDEG-PPIN nodes and list of edges.

Supplementary Figure 1 | The degree distribution of the refined SHIDEG-PPIN.

Data Sheet 1 | The computational scripts plus an example of raw data used in this study.

REFERENCES

Amiri, M., Jafari, M., Azimzadeh Jamalkandi, S., and Davoodi, S.-M. (2013). Atopic dermatitis-associated protein interaction network lead to new insights in chronic sulfur mustard skin lesion mechanisms. *Expert Rev. Proteomics* 10, 449–460. doi: 10.1586/14789450.2013.841548

Ansari-Pour, N., Razaghi-Moghadam, Z., Barneh, F., and Jafari, M. (2016). Testis-specific Y-centric protein-protein interaction network provides clues to the etiology of severe spermatogenic failure. *J. Proteome Res.* 15, 1011–1022. doi: 10.1021/acs.jproteome.5b01080

Azimzadeh Jamalkandi, S., Mirzaie, M., Jafari, M., Mehrani, H., Shariati, P., and Khodabandeh, M. (2015). Signaling network of lipids as a comprehensive scaffold for omics data integration in sputum of COPD patients. *Biochim. Biophys. Acta* 1851, 1383–1393. doi: 10.1016/j.bbalip.2015.07.005

Baloul, L., Camelo, S., and Lafon, M. (2004). Up-regulation of Fas ligand (FasL) in the central nervous system: a mechanism of immune evasion by rabies virus. *J. Neurovirol.* 10, 372–382. doi: 10.1080/13550280490521122

Baloul, L., and Lafon, M. (2003). Apoptosis and rabies virus neuroinvasion. *Biochimie* 85, 777–788. doi: 10.1016/S0300-9084(03)00137-8

Bastian, M., Heymann, S., and Jacomy, M. (2009). "Gephi: an open source software for exploring and manipulating networks," in *International AAAI Conference on Weblogs and Social Media*.

Bender, R. H. F., Haigis, K. M., and Gutmann, D. H. (2015). Activated k-ras, but not h-ras or N-ras, regulates brain neural stem cell proliferation in a raf/rb-dependent manner. *Stem Cells* 33, 1998–2010. doi: 10.1002/stem.1990

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi: 10.1093/bioinformatics/btp101

Birrell, M. A., Bonvini, S. J., Wortley, M. A., Buckley, J., Yew-Booth, L., Maher, S. A., et al. (2015). The role of adenylyl cyclase isoform 6 in beta-adrenoceptor signalling in murine airways. *Br. J. Pharmacol.* 172, 131–141. doi: 10.1111/bph.12905

Bowick, G. C., and McAuley, A. J. (2011). Meta-analysis of high-throughput datasets reveals cellular responses following hemorrhagic fever virus infection. *Viruses* 3, 613–619. doi: 10.3390/v3050613

Brzózka, K., Finke, S., and Conzelmann, K. K. (2006). Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2. *J. Virol.* 80, 2675–2683. doi: 10.1128/JVI.80.6.2675-2683.2006

Camacho-Cáceres, K. I., Acevedo-Díaz, J. C., Pérez-Marty, L. M., Ortiz, M., Irizarry, J., Cabrera-Ríos, M., et al. (2015). Multiple criteria optimization joint analyses of microarray experiments in lung cancer: from existing microarray data to new knowledge. *Cancer Med.* 4, 1884–1900. doi: 10.1002/cam4.540

Ceccaldi, P. E., Valtorta, F., Braud, S., Hellio, R., and Tsiang, H. (1997). Alteration of the actin-based cytoskeleton by rabies virus. *J. Gen. Virol.* 78(Pt 11), 2831–2835. doi: 10.1099/0022-1317-78-11-2831

Chen, X., Liang, S., Zheng, W., Liao, Z., Shang, T., and Ma, W. (2008). Meta-analysis of nasopharyngeal carcinoma microarray data explores mechanism of EBV-regulated neoplastic transformation. *BMC Genomics* 9:322. doi: 10.1186/1471-2164-9-322

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14:128. doi: 10.1186/1471-2105-14-128

Colledge, M., Dean, R. A., Scott, G. K., Langeberg, L. K., Huganir, R. L., and Scott, J. D. (2000). Targeting of PKA to glutamate receptors through a MAGUK-AKAP complex. *Neuron* 27, 107–119. doi: 10.1016/S0896-6273(00)00013-1

Conley, J. M., Brand, C. S., Bogard, A. S., Pratt, E. P., Xu, R., Hockerman, G. H., et al. (2013). Development of a high-throughput screening paradigm for the discovery of small-molecule modulators of adenylyl cyclase: identification of an adenylyl cyclase 2 inhibitor. *J. Pharmacol. Exp. Ther.* 347, 276–287. doi: 10.1124/jpet.113.207449

Costa, R. M., Honjo, T., and Silva, A. J. (2003). Learning and memory deficits in Notch mutant mice. *Curr. Biol.* 13, 1348–1354. doi: 10.1016/S0960-9822(03)00492-5

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., et al. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Res.* 42, D472–D477. doi: 10.1093/nar/gkt1102

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Compl. Syst.* 1695.

Depiereux, S., Le Gac, F., De Meulder, B., Pierre, M., Helaers, R., Guiguen, Y., et al. (2015). Meta-analysis of microarray data of rainbow trout fry gonad differentiation modulated by ethynylestradiol. *PLoS ONE* 10:e0135799. doi: 10.1371/journal.pone.0135799

de Souza, A., and Madhusudana, S. N. (2014). Survival from rabies encephalitis. *J. Neurol. Sci.* 339, 8–14. doi: 10.1016/j.jns.2014.02.013

Dhingra, V., Li, X., Liu, Y., and Fu, Z. F. (2007). Proteomic profiling reveals that rabies virus infection results in differential expression of host proteins involved in ion homeostasis and synaptic physiology in the central nervous system. *J. Neurovirol.* 13, 107–117. doi: 10.1080/13550280601178226

Dorfmeier, C. L., Tzvetkov, E. P., Gatt, A., and McGettigan, J. P. (2013). Investigating the role for IL-21 in rabies virus vaccine-induced immunity. *PLoS Negl. Trop. Dis.* 7:e2129. doi: 10.1371/journal.pntd.0002129

Farahtaj, F., Zandi, F., Khalaj, V., Biglari, P., Fayaz, A., and Vaziri, B. (2013). Proteomics analysis of human brain tissue infected by street rabies virus. *Mol. Biol. Rep.* 40, 6443–6450. doi: 10.1007/s11033-013-2759-0

Faul, E. J., Wanjalla, C. N., Suthar, M. S., Gale, M., Wirblich, C., and Schnell, M. J. (2010). Rabies virus infection induces type I interferon production in an IPS-1 dependent manner while dendritic cell activation relies on IFNAR signaling. *PLoS Pathog.* 6:e1001016. doi: 10.1371/journal.ppat.1001016

Francischetti, I. M., Assumpção, T. C., Ma, D., Li, Y., Vicente, E. C., Uieda, W., et al. (2013). The "Vampirome": transcriptome and proteome analysis of the principal and accessory submaxillary glands of the vampire bat *Desmodus rotundus*, a vector of human rabies. *J. Proteomics* 82, 288–319. doi: 10.1016/j.jprot.2013.01.009

Fu, Z. F., and Jackson, A. C. (2005). Neuronal dysfunction and death in rabies virus infection. *J. Neurovirol.* 11, 101–106. doi: 10.1080/13550280590900445

Fu, Z. F., Li, X., and Dhingra, V. (2008). Pathogenic rabies virus alters host protein expression in the central nervous system: implications for neuronal dysfunction. *Dev. Biol. (Basel).* 131, 83–91.

Galelli, A., Baloul, L., and Lafon, M. (2000). Abortive rabies virus central nervous infection is controlled by T lymphocyte local recruitment and induction of apoptosis. *J. Neurovirol.* 6, 359–372. doi: 10.3109/13550280009018300

Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315. doi: 10.1093/bioinformatics/btg405

Gitter, A., Klein-Seetharaman, J., Gupta, A., and Bar-Joseph, Z. (2011). Discovering pathways by orienting edges in protein interaction networks. *Nucleic Acids Res.* 39, e22. doi: 10.1093/nar/gkq1207

Gomme, E. A., Wirblich, C., Addya, S., Rall, G. F., and Schnell, M. J. (2012). Immune clearance of attenuated rabies virus results in neuronal survival with altered gene expression. *PLoS Pathog.* 8:e1002971. doi: 10.1371/journal.ppat.1002971

Good, P. F., Olanow, C. W., and Perl, D. P. (1992). Neuromelanin-containing neurons of the substantia nigra accumulate iron and aluminum in Parkinson's disease: a LAMMA study. *Brain Res.* 593, 343–346. doi: 10.1016/0006-8993(92)91334-B

Gourmelon, P., Briet, D., Court, L., and Tsiang, H. (1986). Electrophysiological and sleep alterations in experimental mouse rabies. *Brain Res.* 398, 128–140. doi: 10.1016/0006-8993(86)91258-8

Han, M. G., Park, J. S., Lee, C. S., Jeong, Y. E., Ryou, J. S., Cho, J. E., et al. (2011). Serum microRNA expression profiling in mice infected with rabies virus. *Osong. Public Health Res. Perspect.* 2, 186–191. doi: 10.1016/j.phrp.2011.11.043

Hardin, J., and Wilson, J. (2009). A note on oligonucleotide expression values not being normally distributed. *Biostatistics* 10, 446–450. doi: 10.1093/biostatistics/kxp003

Hemachudha, T., Panpanich, T., Phanuphak, P., Manatsathit, S., and Wilde, H. (1993). Immune activation in human rabies. *Trans. R. Soc. Trop. Med. Hyg.* 87, 106–108. doi: 10.1016/0035-9203(93)90446-W

Hua, F., Hautaniemi, S., Yokoo, R., and Lauffenburger, D. A. (2006). Integrated mechanistic and data-driven modelling for multivariate analysis of signalling pathways. *J. R. Soc. Interface* 3, 515–526. doi: 10.1098/rsif.2005.0109

Huang, Y., Jiao, S., Tao, X., Tang, Q., Jiao, W., Xiao, J., et al. (2014). Met-CCL5 represents an immunotherapy strategy to ameliorate rabies virus infection. *J. Neuroinflammation* 11:146. doi: 10.1186/s12974-014-0146-y

Hunter, D. J. (2005). Gene–environment interactions in human diseases. *Nat. Rev. Genet.* 6, 287–298. doi: 10.1038/nrg1578

Ito, N., Mita, T., Shimizu, K., Ito, Y., Masatani, T., Nakagawa, K., et al. (2011). Amino acid substitution at position 95 in rabies virus matrix protein affects viral pathogenicity. *J. Vet. Med. Sci.* 73, 1363–1366. doi: 10.1292/jvms.11-0151

Iwata, M., Komori, S., Unno, T., Minamoto, N., and Ohashi, H. (1999). Modification of membrane currents in mouse neuroblastoma cells following infection with rabies virus. *Br. J. Pharmacol.* 126, 1691–1698. doi: 10.1038/sj.bjp.0702473

Jafari, M., Mirzaie, M., and Sadeghi, M. (2015). Interlog protein network: an evolutionary benchmark of protein interaction networks for the evaluation of clustering algorithms. *BMC Bioinformatics* 16:319. doi: 10.1186/s12859-015-0755-1

Jafari, M., Sadeghi, M., Mirzaie, M., Marashi, S.-A., and Rezaei-Tavirani, M. (2013). Evolutionarily conserved motifs and modules in mitochondrial protein-protein interaction networks. *Mitochondrion* 13, 668–675. doi: 10.1016/j.mito.2013.09.006

Johnson, N., Mansfield, K. L., Hicks, D., Nunez, A., Healy, D. M., Brookes, S. M., et al. (2008). Inflammatory responses in the nervous system of mice infected with a street isolate of rabies virus. *Dev. Biol. (Basel).* 131, 65–72.

Junyent, F., Alvira, D., Yeste-Velasco, M., de la Torre, A. V., Beas-Zarate, C., Sureda, F. X., et al. (2010). Prosurvival role of JAK/STAT and Akt signaling pathways in MPP+-induced apoptosis in neurons. *Neurochem. Int.* 57, 774–782. doi: 10.1016/j.neuint.2010.08.015

Kammouni, W., Wood, H., Saleh, A., Appolinario, C. M., Fernyhough, P., and Jackson, A. C. (2015). Rabies virus phosphoprotein interacts with mitochondrial Complex I and induces mitochondrial dysfunction and oxidative stress. *J. Neurovirol.* 21, 370–382. doi: 10.1007/s13365-015-0320-8

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42, D199–D205. doi: 10.1093/nar/gkt1076

Kasempimolporn, S., Lumlertdacha, B., Chulasugandha, P., Boonchang, S., and Sitprija, V. (2014). Alterations in brain cerebral cortex proteome of rabies-infected cat. *Southeast Asian J. Trop. Med. Public Health* 45, 808–815.

Kelder, T., van Iersel, M. P., Hanspers, K., Kutmon, M., Conklin, B. R., Evelo, C. T., et al. (2012). WikiPathways: building research communities on biological pathways. *Nucl. Acids Res.* 40, D1301–D1307. doi: 10.1093/nar/gkr1074

Kennedy, M. B. (1997). The postsynaptic density at glutamatergic synapses. *Trends Neurosci.* 20, 264–268. doi: 10.1016/S0166-2236(96)01033-8

Kluge, S., Rourou, S., Vester, D., Majoul, S., Benndorf, D., Genzel, Y., et al. (2013). Proteome analysis of virus-host cell interaction: rabies virus replication in Vero cells in two different media. *Appl. Microbiol. Biotechnol.* 97, 5493–5506. doi: 10.1007/s00253-013-4939-1

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucl. Acids Res.*, 44, W90–W97. doi: 10.1093/nar/gkw377

Koschel, K., and Halbach, M. (1979). Rabies virus infection selectively impairs membrane receptor functions in neuronal model cells. *J. Gen. Virol.* 42, 627–632. doi: 10.1099/0022-1317-42-3-627

Koschel, K., and Münzel, P. (1984). Inhibition of opiate receptor-mediated signal transmission by rabies virus in persistently infected NG-108-15 mouse neuroblastoma-rat glioma hybrid cells. *Proc. Natl. Acad. Sci. U.S.A.* 81, 950–954. doi: 10.1073/pnas.81.3.950

Lahaye, X., Vidy, A., Fouquet, B., and Blondel, D. (2012). Hsp70 protein positively regulates rabies virus infection. *J. Virol.* 86, 4743–4751. doi: 10.1128/JVI.06501-11

Lahaye, X., Vidy, A., Pomier, C., Obiang, L., Harper, F., Gaudin, Y., et al. (2009). Functional characterization of Negri bodies (NBs) in rabies virus-infected cells: evidence that NBs are sites of viral transcription and replication. *J. Virol.* 83, 7948–7958. doi: 10.1128/JVI.00554-09

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Larrous, F., Gholami, A., Mouhamad, S., Estaquier, J., and Bourhy, H. (2010). Two overlapping domains of a lyssavirus matrix protein that acts on different cell death pathways. *J. Virol.* 84, 9897–9906. doi: 10.1128/JVI.00761-10

Lay, S., Préhaud, C., Dietzschold, B., and Lafon, M. (2003). Glycoprotein of nonpathogenic rabies viruses is a major inducer of apoptosis in human jurkat T cells. *Ann. N. Y. Acad. Sci.* 1010, 577–581. doi: 10.1196/annals.1299.108

Li, A., and Horvath, S. (2007). Network neighborhood analysis with the multi-node topological overlap measure. *Bioinformatics* 23, 222–231. doi: 10.1093/bioinformatics/btl581

Li, J., Ertel, A., Portocarrero, C., Barkhouse, D. A., Dietzschold, B., Hooper, D. C., et al. (2012). Postexposure treatment with the live-attenuated rabies virus (RV) vaccine TriGAS triggers the clearance of wild-type RV from the Central Nervous System (CNS) through the rapid induction of genes relevant to adaptive immunity in CNS tissues. *J. Virol.* 86, 3200–3210. doi: 10.1128/JVI.06699-11

Lieu, K. G., Brice, A., Wiltzer, L., Hirst, B., Jans, D. A., Blondel, D., et al. (2013). The rabies virus interferon antagonist P protein interacts with activated STAT3 and inhibits Gp130 receptor signaling. *J. Virol.* 87, 8261–8265. doi: 10.1128/JVI.00989-13

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lüscher, C., Nicoll, R. A., Malenka, R. C., and Muller, D. (2000). Synaptic plasticity and dynamic modulation of the postsynaptic membrane. *Nat. Neurosci.* 3, 545–550. doi: 10.1038/75714

McConnachie, G., Langeberg, L. K., and Scott, J. D. (2006). AKAP signaling complexes: getting to the heart of the matter. *Trends Mol. Med.* 12, 317–323. doi: 10.1016/j.molmed.2006.05.008

Megid, J., Appolinario, C. M., Mazzini, A. M., and Almeida, M. F. (2006). Evaluation of cytokines concentration and percentage of survival of rabies virus-infected mice submitted to anti-rabies Vero-cell propagated vaccine and *P. acnes*. *Vet. Immunol. Immunopathol.* 114, 192–196. doi: 10.1016/j.vetimm.2006.07.010

Mehta, S. M., Banerjee, S. M., and Chowdhary, A. S. (2015). Postgenomics biomarkers for rabies-the next decade of proteomics. *OMICS* 19, 67–79. doi: 10.1089/omi.2014.0127

Ménager, P., Roux, P., Mégret, F., Bourgeois, J.-P., Le Sourd, A.-M., Danckaert, A., et al. (2009). Toll-like receptor 3 (TLR3) plays a major role in the formation of rabies virus Negri Bodies. *PLoS Pathog.* 5:e1000315. doi: 10.1371/journal.ppat.1000315

Meslin, F. X., Kaplan, M. M., and Koprowski, H. (1996). *Laboratory Technique in Rabies. World Health Organization.* Geneva: World Health Organization.

Moseley, G. W., Lahaye, X., Roth, D. M., Oksayan, S., Filmer, R. P., Rowe, C. L., et al. (2009). Dual modes of rabies P-protein association with microtubules: a novel strategy to suppress the antiviral response. *J. Cell Sci.* 122(Pt 20), 3652–3662. doi: 10.1242/jcs.045542

Mundschau, L. J., and Faller, D. V. (1994). Endogenous inhibitors of the dsRNA-dependent eIF-2 alpha protein kinase PKR in normal and ras-transformed cells. *Biochimie* 76, 792–800. doi: 10.1016/0300-9084(94)90083-3

Nakamichi, K., Inoue, S., Takasaki, T., Morimoto, K., and Kurane, I. (2004). Rabies virus stimulates nitric oxide production and CXC chemokine ligand 10 expression in macrophages through activation of extracellular signal-regulated kinases 1 and 2. *J. Virol.* 78, 9376–9388. doi: 10.1128/JVI.78.17.9376-9388.2004

Nakamichi, K., Saiki, M., Sawada, M., Takayama-Ito, M., Yamamuro, Y., Morimoto, K., et al. (2005). Rabies virus-induced activation of mitogen-activated protein kinase and NF-kappaB signaling pathways regulates expression of CXC and CC chemokine ligands in microglia. *J. Virol.* 79, 11801–11812. doi: 10.1128/JVI.79.18.11801-11812.2005

Paraboschi, E., Cardamone, G., Rimoldi, V., Gemmati, D., Spreafico, M., Duga, S., et al. (2015). Meta-analysis of multiple sclerosis microarray data reveals dysregulation in RNA splicing regulatory genes. *Int. J. Mol. Sci.* 16, 23463–23481. doi: 10.3390/ijms161023463

Patel, T. B., Du, Z., Pierre, S., Cartin, L., and Scholich, K. (2001). Molecular biological approaches to unravel adenylyl cyclase signaling and function. *Gene* 269, 13–25. doi: 10.1016/S0378-1119(01)00448-6

Préhaud, C., Lay, S., Dietzschold, B., and Lafon, M. (2003). Glycoprotein of nonpathogenic rabies viruses is a key determinant of human cell apoptosis. *J. Virol.* 77, 10537–10547. doi: 10.1128/JVI.77.19.10537-10547.2003

Quaranta, A., Siniscalchi, M., Albrizio, M., Volpe, S., Buonavoglia, C., and Vallortigara, G. (2008). Influence of behavioural lateralization on interleukin-2 and interleukin-6 gene expression in dogs before and after immunization with rabies vaccine. *Behav. Brain Res.* 186, 256–260. doi: 10.1016/j.bbr.2007.08.014

Ramasamy, A., Mondry, A., Holmes, C. C., and Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* 5:e184. doi: 10.1371/journal.pmed.0050184

Rieder, M., Brzózka, K., Pfaller, C. K., Cox, J. H., Stitz, L., and Conzelmann, K. K. (2011). Genetic dissection of interferon-antagonistic functions of rabies virus phosphoprotein: inhibition of interferon regulatory factor 3 activation is important for pathogenicity. *J. Virol.* 85, 842–852. doi: 10.1128/JVI.01427-10

Rupprecht, C. E. (1996). "Rhabdoviruses: rabies virus," in *Medical Microbiology, 4th Edn.*, ed S. Baron (Galveston, TX: University of Texas Medical Branch at Galveston).

Russell, S. J. (2002). RNA viruses as virotherapy agents. *Cancer Gene Ther.* 9, 961–966. doi: 10.1038/sj.cgt.7700535

Sagara, J., Tsukita, S., Yonemura, S., Tsukita, S., and Kawai, A. (1995). Cellular actin-binding ezrin-radixin-moesin (ERM) family proteins are incorporated into the rabies virion and closely associated with viral envelope proteins in the cell. *Virology* 206, 485–494. doi: 10.1016/S0042-6822(95)80064-6

Sarmento, L., Tseggai, T., Dhingra, V., and Fu, Z. F. (2006). Rabies virus-induced apoptosis involves caspase-dependent and caspase-independent pathways. *Virus Res.* 121, 144–151. doi: 10.1016/j.virusres.2006.05.002

Schnell, M. J., McGettigan, J. P., Wirblich, C., and Papaneri, A. (2010). The cell biology of rabies virus: using stealth to reach the brain. *Nat. Rev. Microbiol.* 8, 51–61. doi: 10.1038/nrmicro2260

Schwencke, C., Yamamoto, M., Okumura, S., Toya, Y., Kim, S. J., and Ishikawa, Y. (1999). Compartmentation of cyclic adenosine 3′,5′-monophosphate signaling in caveolae. *Mol. Endocrinol.* 13, 1061–1070. doi: 10.1210/mend.13.7.0304

Sharma, M., Ying, R., Tarr, G., and Barnabas, R. (2015). Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. *Nature* 528, S77–S85. doi: 10.1038/nature16044

Silva, S. R., Katz, I. S., Mori, E., Carnieli, P. Jr., Vieira, L. F., Batista, H. B., et al. (2013). Biotechnology advances: a perspective on the diagnosis and research of Rabies Virus. *Biologicals* 41, 217–223. doi: 10.1016/j.biologicals.2013.04.002

Silverbush, D., and Sharan, R. (2014). Network orientation via shortest paths. *Bioinformatics* 30, 1449–1455. doi: 10.1093/bioinformatics/btu043

Song, G. G., Kim, J.-H., Seo, Y. H., Choi, S. J., Ji, J. D., and Lee, Y. H. (2014). Meta-analysis of differentially expressed genes in primary Sjogren's syndrome by using microarray. *Hum. Immunol.* 75, 98–104. doi: 10.1016/j.humimm.2013.09.012

Song, L., Langfelder, P., and Horvath, S. (2012). Comparison of co-expression measures: mutual information, correlation, and model based indices. *BMC Bioinformatics* 13:328. doi: 10.1186/1471-2105-13-328

Song, Y., Hou, J., Qiao, B., Li, Y., Xu, Y., Duan, M., et al. (2013). Street rabies virus causes dendritic injury and F-actin depolymerization in the hippocampus. *J. Gen. Virol.* 94(Pt 2), 276–283. doi: 10.1099/vir.0.047480-0

Srithayakumar, V., Sribalachandran, H., Rosatte, R., Nadin-Davis, S. A., and Kyle, C. J. (2014). Innate immune responses in raccoons after raccoon rabies virus infection. *J. Gen. Virol.* 95(Pt 1), 16–25. doi: 10.1099/vir.0.053942-0

Sugiura, N., Uda, A., Inoue, S., Kojima, D., Hamamoto, N., Kaku, Y., et al. (2011). Gene expression analysis of host innate immune responses in the central nervous system following lethal CVS-11 infection in mice. *Jpn. J. Infect. Dis.* 64, 463–472.

Sun, J., Hu, J., Luo, D., Markatou, M., Wang, F., Edabollahi, S., et al. (2012). "Combining knowledge and data driven insights for identifying risk factors using electronic health records," in *Annual Symposium Proceedings/AMIA Symposium* (Chicago, IL), 901–910. Available online at: https://www.amia.org/amia2012

Sun, M., Fuentes, S. M., Timani, K., Sun, D., Murphy, C., Lin, Y., et al. (2008). Akt plays a critical role in replication of nonsegmented negative-stranded RNA viruses. *J. Virol.* 82, 105–114. doi: 10.1128/JVI.01520-07

Sunahara, R. K., Dessauer, C. W., and Gilman, A. G. (1996). Complexity and diversity of mammalian adenylyl cyclases. *Annu. Rev. Pharmacol. Toxicol.* 36, 461–480. doi: 10.1146/annurev.pa.36.040196.002333

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerte-Cepas, J., et al. (2014). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucl. Acids Res.* 43(Database issue): D447–D452. doi: 10.1093/nar/gku1003

Tarassishin, L., Suh, H.-S., and Lee, S. C. (2011). Interferon regulatory factor 3 plays an anti-inflammatory role in microglia by activating the PI3K/Akt pathway. *J. Neuroinflamm.* 8, 187–187. doi: 10.1186/1742-2094-8-187

Terada, L. S., and Nwariaku, F. E. (2011). Escaping Anoikis through ROS: ANGPTL4 controls integrin signaling through Nox1. *Cancer Cell* 19, 297–299. doi: 10.1016/j.ccr.2011.02.019

Terrien, E., Chaffotte, A., Lafage, M., Khan, Z., Préhaud, C., Cordier, F., et al. (2012). Interference with the PTEN-MAST2 interaction by a viral protein leads to cellular relocalization of PTEN. *Sci. Signal.* 5:ra58. doi: 10.1126/scisignal.2002941

Thanomsridetchai, N., Singhto, N., Tepsumethanon, V., Shuangshoti, S., Wacharapluesadee, S., Sinchaikul, S., et al. (2011). Comprehensive proteome analysis of hippocampus, brainstem, and spinal cord from paralytic and furious dogs naturally infected with rabies. *J. Proteome Res.* 10, 4911–4924. doi: 10.1021/pr200276u

Thoulouze, M. I., Lafage, M., Montano-Hirose, J. A., and Lafon, M. (1997). Rabies virus infects mouse and human lymphocytes and induces apoptosis. *J. Virol.* 71, 7372–7380.

Thoulouze, M. I., Lafage, M., Yuste, V. J., Baloul, L., Edelman, L., Kroemer, G., et al. (2003a). High level of Bcl-2 counteracts apoptosis mediated by a live rabies virus vaccine strain and induces long-term infection. *Virology* 314, 549–561. doi: 10.1016/S0042-6822(03)00491-4

Thoulouze, M. I., Lafage, M., Yuste, V. J., Kroemer, G., Susin, S. A., Israel, N., et al. (2003b). Apoptosis inversely correlates with rabies virus neurotropism. *Ann. N. Y. Acad. Sci.* 1010, 598–603. doi: 10.1196/annals.1299.112

Tomita, S., Nicoll, R. A., and Bredt, D. S. (2001). PDZ protein interactions regulating glutamate receptor function and plasticity. *J. Cell Biol.* 153, F19–F24. doi: 10.1083/jcb.153.5.f19

Vaziri, B., Torkashvand, F., Eslami, N., and Fayaz, A. (2012). Comparative proteomics analysis of mice lymphocytes in early stages of infection by different strains of rabies virus. *Indian J. Virol.* 23, 311–316. doi: 10.1007/s13337-012-0093-0

Venugopal, A. K., Ghantasala, S. S., Selvan, L. D., Mahadevan, A., Renuse, S., Kumar, P., et al. (2013). Quantitative proteomics for identifying biomarkers for Rabies. *Clin. Proteomics* 10:3. doi: 10.1186/1559-0275-10-3

Vidy, A., Chelbi-Alix, M., and Blondel, D. (2005). Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. *J. Virol.* 79, 14411–14420. doi: 10.1128/JVI.79.22.14411-14420.2005

Wang, C.-Y., Lai, M.-D., Phan, N. N., Sun, Z., and Lin, Y.-C. (2015). Meta-analysis of public microarray datasets reveals voltage-gated calcium gene signatures in clinical cancer patients. *PLoS ONE* 10:e0125766. doi: 10.1371/journal.pone.0125766

Wang, X., Ning, Y., and Guo, X. (2015). Integrative meta-analysis of differentially expressed genes in osteoarthritis using microarray technology. *Mol. Med. Rep.* 12, 3439–3445. doi: 10.3892/mmr.2015.3790

Wang, X., Zhang, S., Sun, C., Yuan, Z. G., Wu, X., Wang, D., et al. (2011). Proteomic profiles of mouse neuro N2a cells infected with variant virulence of rabies viruses. *J. Microbiol. Biotechnol.* 21, 366–373.

Wang, Z. W., Sarmento, L., Wang, Y., Li, X. Q., Dhingra, V., Tseggai, T., et al. (2005). Attenuated rabies virus activates, while pathogenic rabies virus evades, the host innate immune responses in the central nervous system. *J. Virol.* 79, 12554–12565. doi: 10.1128/JVI.79.19.12554-12565.2005

Wong, N., and Wang, X. (2015). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucl. Acids Res.* 43, D146–D152. doi: 10.1093/nar/gku1104

Yin, L., Coelho, S. G., Valencia, J. C., Ebsen, D., Mahns, A., Smuda, C., et al. (2015). Identification of genes expressed in hyperpigmented skin using meta-analysis of microarray datasets. *J. Investig. Dermatol.* 135, 2455–2463. doi: 10.1038/jid.2015.179

Yip, A. M., and Horvath, S. (2007). Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinformatics* 8:22. doi: 10.1186/1471-2105-8-22

Youdim, M. B. H. (2008). Brain iron deficiency and excess; cognitive impairment and neurodegeneration with involvement of striatum and hippocampus. *Neurotoxicol. Res.* 14, 45–56. doi: 10.1007/BF03033574

Yousaf, M. Z., Qasim, M., Zia, S., Khan, M. U. R., Ashfaq, U. A., and Khan, S. (2012). Rabies molecular virology, diagnosis, prevention and treatment. *Virol. J.* 9:50. doi: 10.1186/1743-422X-9-50

Zandi, F., Eslami, N., Soheili, M., Fayaz, A., Gholami, A., and Vaziri, B. (2009). Proteomics analysis of BHK-21 cells infected with a fixed strain of rabies virus. *Proteomics* 9, 2399–2407. doi: 10.1002/pmic.200701007

Zandi, F., Eslami, N., Torkashvand, F., Fayaz, A., Khalaj, V., and Vaziri, B. (2013). Expression changes of cytoskeletal associated proteins in proteomic profiling of neuroblastoma cells infected with different strains of rabies virus. *J. Med. Virol.* 85, 336–347. doi: 10.1002/jmv.23458

Zhao, P., Yang, Y., Feng, H., Zhao, L., Qin, J., Zhang, T., et al. (2013). Global gene expression changes in BV2 microglial cell line during rabies virus infection. *Infect. Genet. Evol.* 20, 257–269. doi: 10.1016/j.meegid.2013.09.016

Zhao, P., Zhao, L., Zhang, K., Feng, H., Wang, H., Wang, T., et al. (2012a). Infection with street strain rabies virus induces modulation of the microRNA profile of the mouse brain. *Virol. J.* 9, 159–159. doi: 10.1186/1743-422X-9-159

Zhao, P., Zhao, L., Zhang, T., Qi, Y., Wang, T., Liu, K., et al. (2011). Innate immune response gene expression profiles in central nervous system of mice infected with rabies virus. *Comp. Immunol. Microbiol. Infect. Dis.* 34, 503–512. doi: 10.1016/j.cimid.2011.09.003

Zhao, P., Zhao, L., Zhang, T., Wang, H., Qin, C., Yang, S., et al. (2012b). Changes in microRNA expression induced by rabies virus infection in mouse brains. *Microb. Pathog.* 52, 47–54. doi: 10.1016/j.micpath.2011.10.001

Ziff, E. B. (1997). Enlightening the postsynaptic density. *Neuron* 19, 1163–1174. doi: 10.1016/S0896-6273(00)80409-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Azimzadeh Jamalkandi, Mozhgani, Gholami Pourbadie, Mirzaie, Noorbakhsh, Vaziri, Gholami, Ansari-Pour and Jafari. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 February 2017
doi: 10.3389/fmicb.2017.00120

Phenotypic Consequences *In vivo* and *In vitro* of Rearranging the P Gene of RABV HEP-Flury

*Mingzhu Mei[1,2], Teng Long[1,2], Qiong Zhang[1,2], Jing Zhao[1,2], Qin Tian[1,2], Jiaojiao Peng[1,2], Jun Luo[1,2], Yifei Wang[1,2], Yingyi Lin[1,2] and Xiaofeng Guo[1,2]**

[1] College of Veterinary Medicine, South China Agricultural University, Guangzhou, China, [2] Key Laboratory of Zoonosis Prevention and Control of Guangdong Province, Guangzhou, China

Phosphoprotein (P) of the Rabies virus (RABV) is critically required for viral replication and pathogenicity. Here we manipulated infectious cDNA clones of the RABV HEP-Flury to translocate the P gene from its wild-type position 2 to 1, 3, or 4 in gene order, using an approach which left the viral nucleotide sequence unaltered. The recovered viruses were evaluated for the levels of gene expression, growth kinetics in cell culture, lethality in suckling mice and protection of mice. The results showed that viral replication was affected by the absolute value of N protein which was regulated by P protein. Viral lethality in suckling mice was consistent with the ratio of P mRNA in one complete transcription. The protection of mice induced by viruses was related to the antibody titer 5 weeks post-infection which might be regulated by G protein. However, the ability to induce cell apoptosis and viral spread were not only related to the viral replication but also to the ratio of related gene which affected by the gene position. These findings might not only improve the understanding of phenotype of RABV and P gene rearrangement, but also help rabies vaccine candidate construction.

Keywords: Rabies virus, HEP-Flury, gene rearrangement, phosphoprotein, pathogenicity

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Takashi Irie,
Hiroshima University, Japan
Guiqing Peng,
Huazhong Agriculture University, China

***Correspondence:**
Xiaofeng Guo
xfguo@scau.edu.cn

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *12 October 2016*
Accepted: *17 January 2017*
Published: *03 February 2017*

Citation:
Mei M, Long T, Zhang Q, Zhao J, Tian Q, Peng J, Luo J, Wang Y, Lin Y and Guo X (2017) Phenotypic Consequences In vivo and In vitro of Rearranging the P Gene of RABV HEP-Flury. Front. Microbiol. 8:120. doi: 10.3389/fmicb.2017.00120

INTRODUCTION

Rabies is a zoonotic disease caused by rabies virus (RABV) belonging to the *Lyssavirus* genus of the *Rhabdoviridae* family, *Mononegavirales* (Pringle, 1996). It is still a major health concern in many developing countries. The RABV genome is approximately 12 kb encoding five proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and the RNA-dependent RNA polymerase (L) (Albertini et al., 2011). RABV P is a multifunctional protein: it is an essential cofactor of the virus RNA-dependent RNA polymerase which is important for the genome transcription/replication, and in addition, it has been identified as an interferon antagonist (Blondel et al., 2002; Brzozka et al., 2005; Vidy et al., 2005; Rieder et al., 2011). The reduced expression of P protein can decrease the ability to prevent IFN induction (Brzozka et al., 2005; Marschalek et al., 2009) and a def-P virus has been demonstrated apathogenic in both adult and suckling mice, even when inoculated intracranial (Shoji et al., 2004; Morimoto et al., 2005). Moreover, wild-type RABV P protein has been reported to assist viral replication in muscle cells by counteracting the host IFN system, consequently, enhancing infection of peripheral nerves (Niu et al., 2013; Yamaoka et al., 2013). Also, P protein can interact with mitochondrial complex I and induces mitochondrial dysfunction and oxidative stress (Kammouni et al., 2015).

Gene rearrangement can alter the genotype of a virus, resulting in a predictable change in gene expression which would be invaluable for studies of gene be invaluable for studies of gene function and control (Wertz et al., 1998). Previous studies have successfully rearranged

the five viral genes of vesicular stomatitis virus (VSV), a prototype *Vesiculovirus* genus of the *Rhabdoviridae* family, and recovered viable viruses from each of the rearranged cDNAs (Wertz et al., 1998; Ball et al., 1999). Subsequently, they demonstrate that neither of the RNA species in variant viruses infecting cells nor the relative molar ratios of the proteins in mature virus particles are changed by gene rearrangement. Gene rearrangement only affects the relative levels of protein expression and consequently alters the phenotypes and lethality in mice infected with recombinant viruses (Wertz et al., 1998; Ball et al., 1999; Flanagan et al., 2000).

Unlike VSV, which is a highly cytopathic virus and should replicate very fast, RABV regulates viral gene expression to produce viral components in sufficient amounts for viral spread, but low enough to maintain host cell survival and to escape from antiviral host cell responses (Albertini et al., 2011). In previous work, G mRNA of RABV ERA can be increased 30% by switching the positions of G gene with M gene (Wu and Rupprecht, 2008). However, the high-egg-passage Flury (HEP-Flury) strain, one of the most attenuated of rabies fixed strains used as a vaccine for humans in Japan, could produce more translatable P and G, less M, and equivalent L mRNA compared to ERA which transcription mode like VSV (Morimoto et al., 1992; Morimoto et al., 2011). And RABV P is important for the viral pathogenicity and antiviral response (Chopy et al., 2011; Niu et al., 2011, 2013; Rieder and Conzelmann, 2011; Fouquet et al., 2015). Therefore, investigation of P gene rearrangement can contribute to development of rabies vaccine.

In previous work, we have rearranged the M gene of RABV HEP-Flury from its wild-type position 3 to 2 or 4 decreasing its replication in BSR cells (Yang et al., 2014). Here, we first rearranged the P gene order of HEP-Flury, translocating it from wild-type position 2 to 1, 3 or 4 and subsequently recovered the infectious viruses. We did this investigating the relationship between gene transcription/expression and viral phenotype caused by P gene rearrangement thus provided an optional approach to rabies vaccine development. The results showed that the viral lethality for suckling mice was in accord with the ratio of P mRNA in one complete transcription which decreased as its gene was moved successively away from promoter proximal position to successive positions down the viral genome. Importantly, these changes occurred leaving the protection of mice intact or better, suggesting that this approach may provide a rational method to achieve a measured and stable degree of attenuation of this type of virus. Moreover, since the *Mononegavirales* have not been observed to undergo homologous recombination, gene rearrangement should be irreversible (Pringle, 1982).

MATERIALS AND METHODS

Mice

Specific pathogen-free (SPF) female Kunming adult and pregnant mice were purchased from Center for Laboratory Animal Science, the Southern Medical University (Guangzhou, China). They were housed under specific-pathogen-free conditions in biosafety level containment in the Laboratory Animal Center of South China Agricultural University. All procedures involving animals and their care were conducted in conformity with NIH guidelines (Care and Animals, 2011) and approved by the Animal Care and Use Committee of the South China Agricultural University.

Viruses and Cells

Recombinant RABV rHEP-Flury and P3 with the P gene in position 3 was previously rescued in our laboratory via reverse genetics (Yang et al., 2014). Baby hamster kidney cells (BHK-21) were used for virus recovery from cDNA and cultured in Dulbecco modified essential medium (DMEM) (Gibco, China), supplemented with 10% fetal bovine serum (FBS) (Gibco, Australia). Mouse neuroblastoma NA cells which were used to amplify recombinant viruses and subsequent experiments, were grown in RPMI 1640 medium (Gibco, China) supplemented with 10% FBS.

Plasmid Construction and Recovery of Recombinant Viruses

The plasmid pHEP-3.0 containing the full-length genomic cDNA of HEP-Flury and four helper plasmids pH-N, pH-P, pH-G, and pH-L were a kind gift from Dr. Kinjiro Morimoto. For detailed information about the plasmids, please refer to Inoue et al. (2003).

Construction of a full-length cDNA clone of the P gene rearranged genome and recovery of infectious viruses has been described previously (Ghanem et al., 2012; Wang et al., 2014; Yang et al., 2014; Luo et al., 2016). All the genes were rearranged from the beginning of the transcription start site (AACA) to the transcription end signal (the poly A signal, AAAAAAA). To rearrange the P gene of RABV without introducing any additional changes into the viral genome, we used inverse PCR to amplify the linearized vector for the P1 plasmid with the P gene in position 1 via primers 1–5′VR (5′-*ACATTTTTGCTTTGC*AACTGACGATGTC-3′; the 15–20 bp homologous sequences are underlined) and 1–3′ VF (5'–*AGGCAACACCACTAA*TAAAATGAAC–3′; the 15-20 bp homologous sequences are underlined). The linearized vector for the P4 plasmid with the P gene in position 4 was synthesized using primers 4–5′VR (5′-*AGTTTTTTTCATGATG*GATATACACAATC-3′; the 15–20 bp homologous sequences are underlined) and 4–3′VF (5′-*TGTATACCAAAAGAAC*AACTAACAACAC-3′; the 15–20 bp homologous sequences are underlined). The primers for the amplification of genes are list in **Table 1**. To avoid mutation, a Phusion High-Fidelity DNA Polymerase (Thermo Scientific, USA) was used following the manufacturer's instructions. Then an efficient homologous-recombinant-based ClonExpress^{TM} MultiS one step cloning method was adopted according to manufacturer's instructions (Vazyme Biotech, Nanjing, China). After plasmid sequencing, the plasmids of the rearranged cDNAs and the four helper plasmids pH-N, pH-P, pH-G, and pH-L were used to co-transfect BHK21 cells via the SuperFect Transfection Reagent (Qiagen, USA) according to manufacturer's instructions. Twelve days later, we collected the supernatants of transfected cells and examined the existence of the rescued virus via direct

TABLE 1 | Oligonucleotides used for construction of RABV genome cDNAs[a].

Primer name	Sequence (5′–3′)	Use
1-P-F	GCAAAGCAAAAATGTAACACTCCTCCTTTCGAACC	Construction for P1 plasmid
1-P-R	ATTGTAGGGGTGTTAGTTTTTTTCATATCGACTCC	
1-N-F	TAACACCCCTACAATGGATGCCGAC	
1-N-R	TTAGTGGTGTTGCCTGTTTTTTTCATGATGGATATAC	
4-M-F	ATCATGAAAAAAACTAACACCACTAATAAAATGAAC	Construction for P4 plasmid
4-M-R	TGAGGGATGTTGCCTGTTTTTTTCACATCCAAGAGGC	
4-G-F	AGGCAACATCCCTCAAAAGACTTAAGG	
4-G-R	GGAGGAGTGTTAATAGTTTTTTTCTCGACTGAAATG	
4-P-F	TATTAACACTCCTCCTTTCGAACCATCC	
4-P-R	GTTCTTTTGGTATACAGTTTTTTTCATATCGACTCCATG	

[a]*15–20 bp homologous sequences are underlined.*

immunofluorescence assay (IFA). Subsequently, the viruses rescued successfully were passaged in NA cells.

Following eight passages in NA cell culture, the gene order of the recovered viruses was determined via reverse transcription (RT)-PCR using three pairs of primers: DF (5′-*CTTAACAACAAAACCAAAGAAGAAGCA*-3′) and PR (5′-*CAT CTCAAGATCGGCCAGACCG*-3′); DF and NR (5′-*TGAAGTT CGGTATAGTACTCC*-3′); and DF and MR (5′-*GTCCTCA TCCCTACAGTTTTTC*-3′). Subsequently, the PCR fragments were sequenced directly.

Lethality in Suckling Mice

The lethality of individual virus was measured in suckling Kunming mice aged 1-day to 3-days, obtained from the Southern Medical University of China. Groups of twelve mice were intracranial (IC) inoculated either with 20ul diluent RPMI 1640 or with serial 10-fold dilutions of individual virus, and then observed daily. The titers of viruses were diluted to $10^{5.5}$FFU/ml before serial dilutions and any mouse dying within 4 days post-inoculation was ignored. The LD_{50} for each virus was calculated via the method of Reed and Muench.

Protection of Mice

Groups of 10 mice aged 6-weeks to 8-weeks were immunized once intramuscularly (IM) with different doses of either rHEP-Flury or one of the variant RABVs. To determine antibody levels, blood samples were collected 21 days post immunization. Serum samples were pooled and heated to 57°C for 30 min to inactivate complement. Mice were then challenged IC with 50 LD_{50} of Challenge Virus Standard (CVS-24) and observed daily for a 28 day period. Survivor numbers were recorded and any mouse dying within 4 days post-challenge was ignored.

Monitoring Antibody Levels in Mice

Groups of five mice aged 6-weeks to 8-weeks old were inoculated IM with 10^5 FFU individual viruses. Subsequent to virus inoculation, blood was collected at weekly intervals. Serum samples were pooled and heated to 57°C for 30 min to inactivate complement. The serum antibody titers were monitored using a Serelisa® Rabies Ab Mono Indirect kit (Synbiotics, France) following manufacturer's introductions. For calculated titer >0.6EU/ml, the animal is considered as protected.

One-Step and Multi-Step Growth Analyses of Viruses in NA Cells

NA cell monolayers were infected with rHep-Flury and variant viruses at a multiplicity of infection (MOI) of 0.01 for multi-step growth curves and a MOI of 3 for one-step growth curves. After 1h of adsorption at 37°C, the inoculum was removed, cells were washed with phosphate buffered saline (PBS) Thermo scientific, China) twice, and 5 ml of fresh PRMI medium containing 5% FBS was added and incubated at 34°C. Samples were harvested at indicated intervals over a 120 h period, and viral titers were quantified via direct fluorescent antibody test (FAT) as described previously described on NA monolayers (Zhao et al., 2009).

Viral Spread in NA Cells

NA cell monolayers were infected with rHep-Flury and variant viruses at a MOI of 0.005, then incubated for a 72h period at 34°C, and stained every 12 h with FITC Anti-Rabies Monoclonal Globulin (Fujirebio, Malvern, PA, USA), before they were examined under a fluorescence microscope(Wirblich and Schnell, 2011).

Cell Apoptosis by Flow Cytometry

Cell apoptosis was quantified by using an Annexin V-FITC apoptosis kit (BestBio, China) according to the manufacturer's instructions. NA cells were seeded into 6-well plates and incubated at 37°C overnight. Then cells were treated with RABV rHEP-Flury and P gene rearranged viruses at a MOI of 3. Twenty-four hours later, cells were collected and incubated with 5 μl Annexin V-FITC and 10 μl PI for 15 min. Finally, 500 μl of binding buffer was added to each tube and analyzed by a Beckman FC 500 flow cytometry(Beckman Coulter, Fullerton, CA, USA), followed by data analysis with the corresponding CXP Software.

RNA Isolation and qRT-PCR

A MOI of 3 was chosen to make sure every cell was infected. Monolayer of NA cells grown in six-well plates were infected

with rHep-Flury and variant viruses, respectively, and incubated at 34° C for 12 h. Cells were then washed once with PBS, and RNA was isolated using HiPure Universal RNA Kits (Magen, Guangzhou, China) following the manufacturer's instructions at the indicated intervals. For the viral structural gene expression, cDNAs were synthesized with oligo/ (dT_{23}) primer using the HiScript® II 1st Strand cDNA Synthesis Kit (Vazyme Biotech, Nanjing, China). For the quantification of leader RNA, cDNA was synthesized with a tagged primer with attached 18-nucleotide (nt) tag that was unrelated to RABV as previous described (Yang et al., 2015a). For the quantification of genomic RNA (vRNA), cDNA was synthesized with N-QF. **Table 2** provides primer sequence details. The real-time SYBR Green PCR assay was carried out in a CFX384 Real-time System (Bio-Rad, USA) using Universal SYBR Green Master (Vazyme Biotech, Nanjing, China) according to the manufacturer's instructions. Numbers of RNAs copies of a particular gene were normalized in relation to the housekeeping gene beta actin (β-actin).

Analysis of Viral Protein Synthesis by Western Blotting

Monolayer of NA cells cultured in six-well plates were infected with rHep-Flury and variant viruses at a MOI of 3 and incubated at 34°C for RNAs analysis. At 12 post-infection, cells were washed in PBS once and lysed with RIPA buffer (containing 1× protease inhibitor cocktail) (Beyotime Biotech, China) on ice for 30 min. The suspension was then transferred to a microcentrifuge tube and spun for 20 min at 15,000 × *g* to remove all cell debris, before the suspension was quantified using a Pierce BCA Protein Assay Kit (Thermo scientific, USA). Proteins were separated by SDS-10% polyacrylamide gel electrophoresis (SDS-10% PAGE) and then transferred onto a polyvinylidene difluoride (PVDF) membrane (Millipore, USA). Blots were blocked in 5% dry milk powder in PBS for 1h. After blocking, blots were washed twice with a 0.1% PBS-Tween 20 solution and incubated overnight at 4°C with monoclonal mouse anti-RV N (Tongdian Biotech, Hangzhou, China) (diluted 1:1000), P (prepared in our lab) (diluted 1:500), M (prepared in our lab) (diluted 1:100) or G (prepared in our lab) (diluted 1:500), respectively. The β-actin (1:1000, Beyotime Biotech, China) was as the reference protein. Subsequently, blots were then washed four times with 0.1% PBS-Tween 20. Secondary goat anti-mouse horseradish peroxidase-conjugated antibodies (Bioworld Technology, USA) (diluted 1:50,000) were added, and blots were incubated for 2 h at 37°C. Blots were washed four times with 0.1% PBS-Tween 20 and once with PBS. Chemiluminescence analysis using BeyoECL plus (Beyotime Biotech, China) was performed as instructed by the vendor.

Statistical Analysis

All results were expressed as the mean ± standard deviation (SD) and all statistical analyses were performed with one-way or two-way analysis of variance (ANOVA). Asterisks denote statistical differences ($^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$) between different groups. A *P*-value of less than 0.05 was considered statistically significant. The statistical significance of survival rates was determined by the log-rank test and Kaplan–Meier survival analysis.

RESULTS

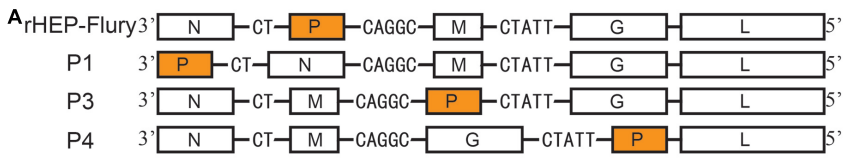
Recovery of Rearranged Viruses

We rearranged the P gene of HEP-Flury by manipulating an infectious cDNA clone to translocate it from its normal position 2 to 1, 3, or 4 in gene order and rescued them successfully (**Figure 1A**). All other aspects of the viral nucleotide sequences remained unaltered.

The gene orders for each of the recovered viruses were verified after eight passages by RT-PCR carried out using three pairs of primers. The observed sizes of the amplified products were exactly as predicted (**Figure 1B**) and further direct sequencing demonstrated that they were specific bands. These

TABLE 2 | Oligonucleotides used for quantification of RABV structural gene and leader RNA.

Primer name	Sequence (5′–3′)	Use
N-QF	TAGGCTTGAGTGGGAAGTC	qRT-PCR/cDNA synthesis
N-QR	CAGCAATAACCGTGGCAT	qRT-PCR
P-QF	ATCGCTCATCAGATTGCT	qRT-PCR
P-QR	GCCTCTTTAACTATGTCATCAA	qRT-PCR
M-QF	GAACATACGGGCTTAACTCC	qRT-PCR
M-QR	AAGAGGCTCAAAATGTAACGG	qRT-PCR
G-QF	TCTACAGTTTTCAAAGACGG	qRT-PCR
G-QR	TCAATACATACTTCCCCCAT	qRT-PCR
L-QF	TGTTGATGTCTGATTTCGC	qRT-PCR
L-QR	GGAACGCTCTTGACAGAT	qRT-PCR
β-ACTIN-QF	CGTAAAGACCTCTATGCCAAC	qRT-PCR
β-ACTIN-QR	GATCTTGATCTTCATGGTGCT	qRT-PCR
LRNA-QF	CCAGATGCTTGGCGTCCT	qRT-PCR
LRNA-QR	ACGCTTAACAACAAAACC	qRT-PCR
LRNA RT	CCAGATGCTTGGCGTCCTCTTTGCAACTGACGATGT	cDNA synthesis



B

Gene order	D-F and P-R(bp)	D-F and N-R(bp)	D-F and M-R(bp)
rHEP-Flury	1570	677	2537
P1	144	1670	2537
P3	2380	677	1541
P4	4141	677	1541

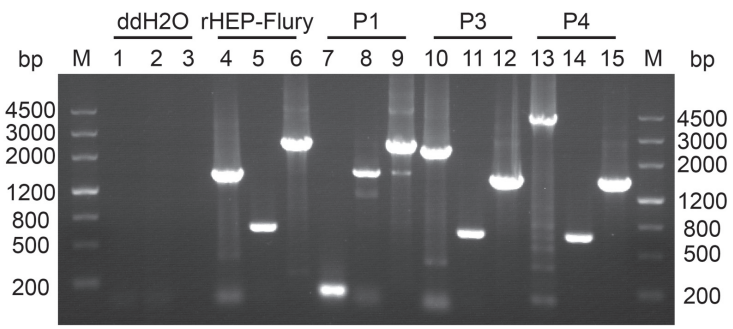


FIGURE 1 | Determination of gene order in recovered viral genomes. (A) Schematic representations of the gene order of rHep-Flury and the rearranged variant viruses. **(B)** Genomic RNA was isolated from recovered viruses and analyzed by reverse transcription and PCR with different primer pair. Fragments were analyzed by electrophoresis on a 1% agarose gel in the presence of ethidium bromide. Lane M = marker DNA fragments with sizes as indicated. Lanes 1, 4, 7, 10, and 13: PCR products with primers D–F and P–R; Lanes 2, 5, 8, 11, and 14: PCR products with primers D–F and N–R; Lanes 3, 6, 9, 12, and 15: PCR products with primers D–F and M–R.

data indicated that gene orders of the recovered viruses were as originally constructed and remained so after eight passages in cell culture.

Lethality in Suckling Mice

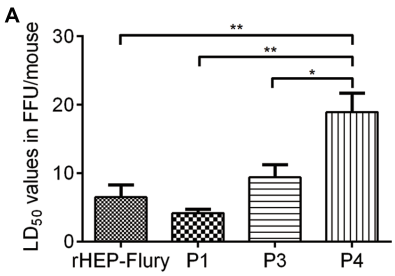
Pathogenicity of recombinant RABVs were assessed in suckling mice as HEP-Flury was fatal for suckling mice but not for the adult mice following IC inoculation (Takayama-Ito et al., 2006). Suckling mice aged 1-day to 3-days served as a sensitive model to compare the relative lethality of rHEP-Flury and its mutants. By IC inoculation, the LD_{50} dose of P4 was significant higher than others. Moreover, the LD_{50} dose of rHEP-Flury increased by 1.7-fold, P3 increased by 2.3-fold, and P4 increased by 4.6-fold compared to P1 (**Figure 2A**). This declared that the LD_{50} dose was going to be higher as the P gene was moved successively away from promoter proximal position to successive positions down the viral genome.

The time to onset of death at doses of $10^{4.5}$ FFU/ml to $10^{2.5}$ FFU/ml per mouse are shown in **Figure 2B**. The rHEP-Flury-infected mice first appeared death at day 5 post-inoculation. Recombinant P3 elicited reproducibly pathogenesis as fast as rHEP-Flury-infected animals, whereas the onset of death from infection with P1 and P4 occurred later as the rule was more clearly with decreasing dose.

Ability of Rearranged Viruses to Protect Against Wild-Type Challenge

To test whether the P gene translocation affected the ability to elicit a protective immune response, mice were immunized by IM inoculation with 10^5 or 10^4FFU of either rHEP-Flury or the variant viruses. The surviving animals were challenged 21 days later by IC inoculation with $50LD_{50}$ of CVS-24. The protections of all viruses were significantly higher than control groups and there was no significant difference between them (**Figure 3A**). We guessed that the P gene rearranged viruses all contained the wild type complement of genes which could induce a protective host response (Wertz et al., 1998).

At a dose of 10^5FFU, P gene rearranged viruses showed the same protections as rHEP-Flury (**Figure 3A**). Consistent with this, there was no significant difference in the serum antibody titer between variant viruses and rHEP-Flury in the immunized animals prior to challenge on day 21 (**Figure 3B**). At a dose of 10^4FFU, the antibody titers of viruses were consistent with the tendency of vital titer at a MOI of 0.01. This suggested that the antibody production was affected by the viral replication. Moreover, the survival rates of P1 and P4 were 100% little better than P3 (77.78%) or rHEP-Flury (87.5%), though their antibody titers were reverse at the dose of 10^4FFU (**Figures 3A,B**). These data revealed



B

Gene order	Days to death(FFU/ml)		
	$10^{4.5}$	$10^{3.5}$	$10^{2.5}$
rHEP-Flury(NPMGL)	5	6	7
P1(PNMGL)	8	9	11
P3(NMPGL)	5	7	9
P4(NMGPL)	7	8	9

FIGURE 2 | Lethality of rHEP-Flury or rearranged viruses for suckling mice. (A) The LD_{50} was calculated from mortality among groups of 12 mice inoculated IC with five serial, 10-fold dilutions of virus. The LD_{50} was analyzed for statistical significance by one-way ANOVA. Asterisks denote statistical differences (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$) between different groups. **(B)** Onset of death after intracerebral inoculation of three serials.

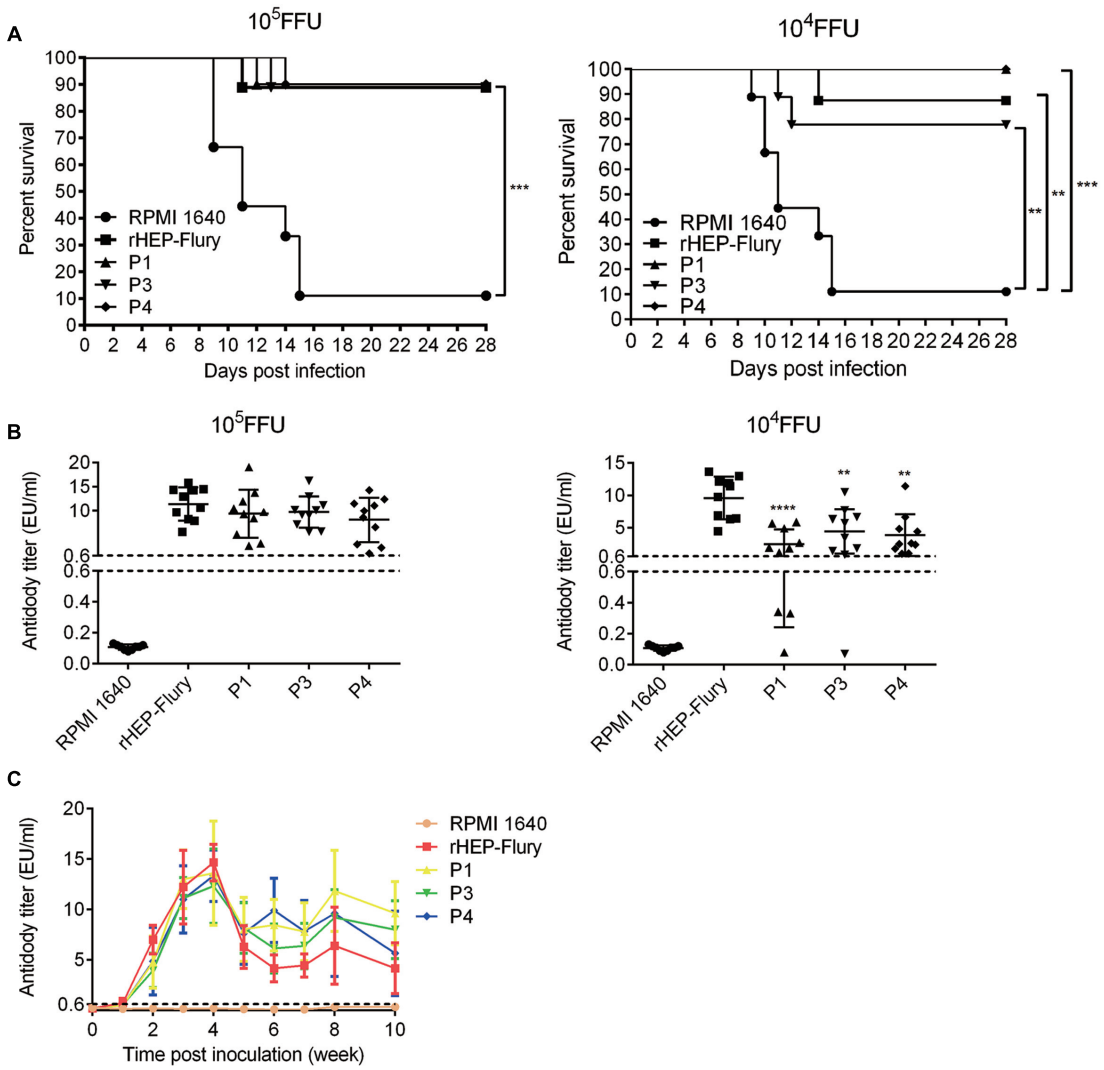


FIGURE 3 | Comparison of antibody production and ability to protect against lethal challenge. (A) Survivorship of mice after immunized with rHEP-Flury, P1, P3, P4, or medium by IM route. The survival rates were analyzed for statistical significance by Kaplan–Meier plots ($n = 10$ in each groups; by log-rank test). **(B)** The presence of anti-RABV antibody in serum at 21 days post-inoculation prior to challenge. The serum antibody titer of P1, P3, or P4 was compared to rHEP-Flury using ordinary one-way ANOVA. Asterisks denote statistical differences (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$) between different groups. **(C)** Antibody kinetics of viruses with P gene translocations. Groups of five mice were inoculated with 10^5 FFU of each virus via IM injection. Serum samples were pooled and heated to 57°C for 30 min to inactivate complement. The serum antibody titers were monitored using a Serelisa® Rabies Ab Mono Indirect kit.

that P gene rearranged viruses elicited a protective response that remained undiminished compared to that of the parent virus.

Duration of Antibody Levels in Mice

Duration of antibody levels is essential for rabies vaccine when a single immunization is conducted. The antibody levels of the serum revealed similar antibody kinetics for all viruses. They reached a maximum at week 4 and then decreased and remained a level more than 0.6EU/ml during 10 weeks post-infection (**Figure 3C**). However, P gene rearranged RABVs showed higher antibody levels than rHEP-Flury 5 weeks post immunization. We speculated that the antibody titers fell evidently at week 5 as the viruses clearance by antibody occurred mainly during this time. And the antibody titer 5 weeks post immunization revealed the final balance between viral replication and antibody development *in vivo*.

Effects of P Gene Rearrangement on Viral Replication and Spread

Slower replication and faster spread could enhance the RABV pathogenicity (Faber et al., 2005; Davis et al., 2015). We investigated viral replication and spread caused by P gene rearrangement in NA cells to further illustrate changes of lethality and immune response. Analysis of progeny virus production in cell culture revealed a decreasing ability to replicate at a MOI of 0.01 due to P gene translocation from its wild-type position 2 to 3 or 4 as predicted (**Figure 4A**). P1, which N gene was in position 2, had the worst ability to replicate and its maximum titer was significant lower than rHEP-Flury. At a MOI of 3, the P gene position affected the speed of growth at early stage. There was no significant difference in maximum titers between them (**Figure 4B**). However, the maximum titer of rHEP-Flury at a MOI of 0.01 was higher than that at a MOI of 3. Viral spread in NA cells also varied as grew at a MOI of 0.01(**Figure 4C**).

Cell Apoptosis

Rabies virus HEP-Flury could induce NA cell apoptosis though RABV dose not induce a typical CPE in NA cells. Cell apoptosis is a particular factor attenuating the pathogenic potential of RABV (Préhaud et al., 2003; Thoulouze et al., 2003; Kassis et al., 2004). Previous work has shown that, at a MOI of 0.01, rHEP-Flury does not cause toxicity in NA cells, and only induced 2.9% NA cell apoptosis at 48 h post-infection (Yang et al., 2015b; Peng et al., 2016). In this study, at a MOI of 3, the percentage of early stage apoptotic cells as well as late stage apoptotic or even necrotic cells was about 11.8%. And P4 induced similar NA cell apoptosis as rHEP-Flury. They were significant more than that induced by P1 or P3 (**Figure 5**). We found that when G gene was in the same position, NA cell apoptosis was positive to the viral replication. The P4 induced more cell apoptosis as the G gene was moved one position closer to the promoter. This indicated that both the viral replication and G gene position were related to cell apoptosis induced by HEP-Flury.

Effects of P Gene Rearrangement on Expression of RNAs and Proteins

Rabies virus is a neurotropic virus (Ugolini, 2011). We analyzed the synthesis of viral RNAs and proteins in infected NA cells to ascertain how P gene translocation affected viral gene expression, thus influencing the phenotype of RABV. Twelve hours post-infection, we stained the NA cells with FITC Anti-Rabies Monoclonal Globulin to confirm the infection of every cell.

RNA and protein profiles of cells infected with the rHEP-Flury and variant viruses showed that both the RNAs and protein levels of rHep-flury were the most at 12 h post-infection (**Figure 6**). That was rHep-flury with the wild-type gene order was always the most fit for growth and gene expression first had to meet the requirement of sufficient virus replication (Finke and Conzelmann, 2005). P mRNA substantially decreased as its gene was moved successively away from the promoter in viruses P1, P3, and P4. The transcription levels of N, M, G and L mRNA were reduced as the vRNA reduced presumably as a secondary effect because of the decrease in replication (**Figure 6A**). Correlation analysis revealed they had significant correlation (correlation coefficients >0.9, $P < 0.001$). RABV only encoded five subgenomic mRNAs that were translated to yield five proteins, all of which were components of the mature virion (Okumura and Harty, 2011b). And at this time, the viruses did not budding from the cells (**Figure 4A**), so they had tightly relationship. However, at 24 and 48 h post-infection, when the viruses budding from the cells which we could found in one-step curve, vRNA levels were on behalf of the balance between synthesis and budding while the level of N, P, M, G, and L mRNA only showed the synthesis, so there was no significant correlation between them (**Supplementary Figure S1**).

Then we analyzed the gene ratio in one complete transcription for each virus, i.e., the ratios of viral structural RNAs were calculated relative to all structural genes plus leader RNA: leader RNA+N mRNA+ P mRNA+ M mRNA+ G mRNA+ L mRNA in every virus. The data showed that the ratio of P mRNA decreased as its gene was moved successively away from the promoter in viruses P1, rHEP-Flury (P2), P3, and P4 (**Figure 6B**). Consistent with this decrease, a decrease in the ratio of N mRNA was also observed with virus P1 in which the N gene was moved one position farther to the promoter; an increase in the ratio of M mRNA was observed with virus P3 or P4 in which the M gene was moved closer to the promoter; an increase in the ratio of G mRNA was observed with virus P4 in which the G gene was moved closer to the promoter (**Figure 6B**). These were as predicted by the model of progressive transcriptional attenuation though previous work has shown that N mRNA and P mRNA of HEP-Flury were qualitatively similar in infected BHK cells (Morimoto et al., 2011). We speculated that gene transcription was also associated with the gene. Moreover, the ratios of leader RNA and L mRNA increased as P gene translocated though their positions in gene order without change. This was an interesting and important observation.

The amounts of viral proteins were qualitatively similar with their mRNAs as the translation efficiency was mainly regulated by the level of transcription (**Figure 6C**). However, the N and

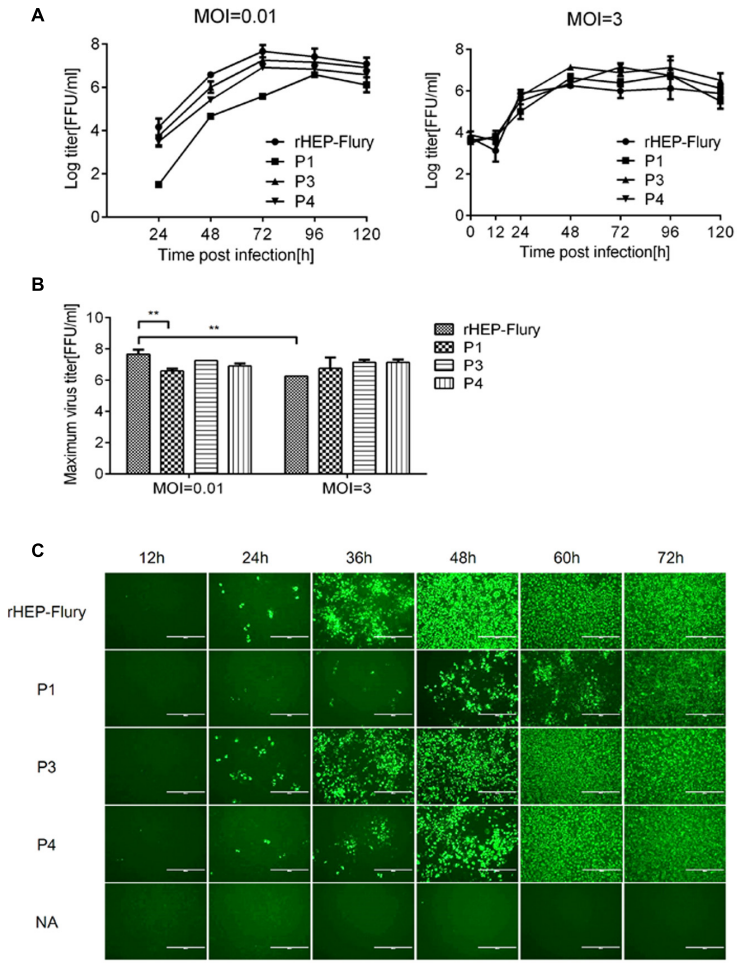


FIGURE 4 | Growth kinetics and spread of RABV strains in NA cells. (A) Multi-step and one-step growth curves of the rHEP-Flury and recombinant strains were shown. NA cells were infected at a MOI of 3 (one-step) or 0.01 (multi-step) and incubated at 34°C for 120 h. Titers of virus harvested at the indicated intervals were assayed in triplicate, and the results were averaged. **(B)** The maximal titers of RABV strains in NA cells at a MOI of 0.01 or 3. The maximal titers were analyzed for statistical significance by two-way ANOVA. Asterisks denote statistical differences (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$). **(C)** Viral spread in NA cells. NA cells were infected with every virus at a MOI of 0.005 for 72 h at 34°C, stained with FITC-conjugated monoclonal antibodies against RABV nucleoprotein, and examined under a fluorescence microscope. For each virus, three representative images taken from different areas of the same well and one of them were shown.

M translation efficiency of P1 were lower than others. There might be some other factors inhibiting the translation when the N gene was translocated from position 1 to position 2, following the P gene translocation. Moreover, G protein levels in NA cells infected by P gene rearranged viruses were more abundant than that infected by rHEP-Flury (**Supplementary Figure S1**). This might induced more effective virus neutralizing antibody (VNA), thus increased the protection of mice.

DISCUSSION

The results presented above revealed that the RABV P gene can be rearranged from its wild type position 2 to 1, 3, and 4 in the genome, leading to successful recovery of infectious viruses. Subsequently, the data showed that gene orders of the recovered viruses correspond to the cDNA clones from which they were recovered. There was no evidence of reappearance of the wild-type genetic order among the variants. As a consequence, it further proved that gene rearrangement should be viable and irreversible (Pringle, 1982). However, the *in vivo*- and *in vitro*-characteristics of rearranged RABVs were different though they could be rescued successfully.

P protein of RABV could interrupt the IFN transcription, consequently increased viral pathogenicity (Kuang et al., 2009; Niu et al., 2013). As our results shown, the viral lethality for suckling mice was in accord with the ratio of P mRNA in one complete transcription which decreased as its gene was moved successively away from promoter proximal position to successive positions down the viral genome, though the absolute value of P mRNA or protein was not consistent with this. Meanwhile, we found that the RABV P4 which G mRNA ratio increased, its

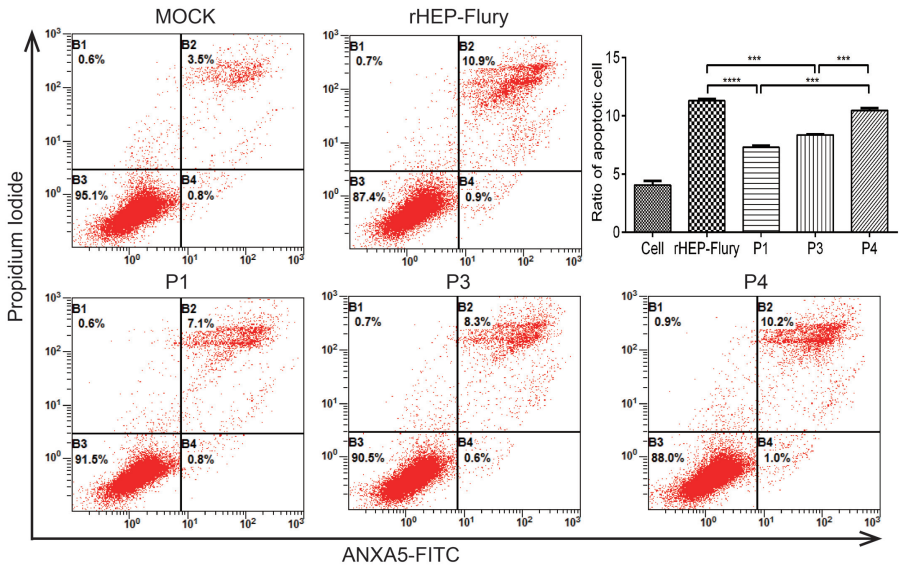


FIGURE 5 | Cell apoptosis induced by rHEP-Flury and rearranged viruses. NA cells were infected with rHEP-Flury and rearranged viruses at a MOI of 3. Cell apoptosis was quantified by using Annexin V-FITC apoptosis kit. Cells in early apoptosis and dead cells were represented as the percentage of ANXA5-FITC and PtdIns cells of total cells. Mean ± SD of three independent experiments. One-way ANOVA: ***$P < 0.001$; ****$P < 0.0001$.

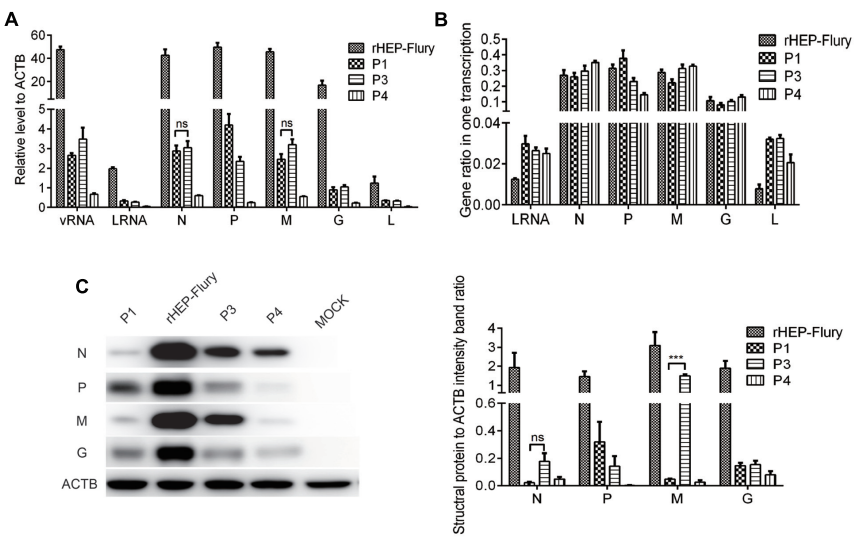


FIGURE 6 | Gene expression in NA cells. NA cells were infected with rHEP-Flury, P1, P3 or P4 at a MOI of 3 for 12 h at 34°C. **(A)** qRT-PCR analyses of RNAs level. The relative amount of individual viral RNAs was normalized in relation to the housekeeping gene β-actin. Data are mean ± SD, $n = 3$. **(B)** Ratios of leader RNA, N mRNA, P mRNA, M mRNA, G mRNA, or L mRNA in one complete transcription. Individual RNA ratio were calculated relative to all structural genes plus leader RNA: leader RNA+N+P+M+G+L. Data are mean ± SD, $n = 3$. **(C)** Viral structural protein were quantified by Western blotting with monoclonal antibody against RABV N, P, M, G, and a monoclonal antibody against actin. Densitometry of the western blotting was analyzed with the Image-Pro Plus 6.0 software.

pathogenicity was significant weaker than others. As HEP-Flury G protein could induce cell apoptosis which was contribute to attenuate the pathogenicity of RABV (Yang et al., 2015b; Peng et al., 2016). We speculated that the pathogenicity of RABV was correlated with the ratio of viral gene. Moreover, fast spread was conceived to viral escaping from antiviral host cell response, thus enhancing the RABV pathogenicity (Faber et al., 2005; Takayama-Ito et al., 2006). Here spread speed mainly affected the onset time of death in suckling mice. We suggested that the viruses have kept away from the majority of the host immune response when inoculated to suckling mice by IC.

Serum antibody levels at 21 days post-inoculation induced by P gene rearranged viruses were lower than rHEP-Flury as their replication efficiency reduced. However, the protections of P1 and P4 were little better than rHEP-Flury and P3 while their antibody titers were reverse. Duration of antibody levels in mice

showed that P gene rearranged RABVs had higher antibody levels than rHEP-Flury 5 weeks post-immunization which might be the effective antibody used to remove the RABV CVS-24. Meanwhile, we found that the antibody levels were consistent with the G protein levels in infected NA cells which could induce VNA (Faber et al., 2002; Li et al., 2006). Moreover, the leader RNA ratio of P gene rearranged viruses increased which was contribute to activate dendritic cells (Kammouni et al., 2015), thus enhancing the protection of mice.

To further illustrate impacts on replication, spread and cell apoptosis by P gene rearrangement, we evaluated them in NA cells. Firstly, we found that vRNA replication varied as their N mRNA levels, though P protein was a noncatalytic cofactor for the polymerase L and conferring the specificity of genomic RNA encapsidation by N (Emerson and Yu, 1975; Liu et al., 2004). This declared N protein played the most important role in viral replication (Albertini et al., 2011; Choi et al., 2015). Though N genes of rHEP-Flury, P3 and P4 were all in position 1, the N gene expression decreased as the P gene was translocated from position 2 to 3 or 4. We suggested optimal N: P: L ratio also had to achieve optimal RNA replication in RABV which has been demonstrated in VSV (Pattnaik and Wertz, 1990). And the P gene rearranged viruses regulated the N gene expression in order to reach the optimal N: P: L to facilitate the viral replication. This might be caused by regulating the binding between N protein and nascent leader RNA (Albertini et al., 2011). The structural M protein of RABV which was an essential factor for virus budding was not only a regulatory protein adjusting the balance of RNP replication and mRNA synthesis but also regulating the viral replication start. (Finke and Conzelmann, 2003; Finke et al., 2003; Davis et al., 2015). Consistent with this work, we found that the replication of rHEP-Flury and P3 were faster than P1 or P4 as their M protein levels were also higher at 12 h post-infection. As for P1, the reduction in viral replication was not only due to the vRNA synthesis but also the inhibition of virion release.

G protein was a key element for viral spread in CNS and a G gene deletion-mutant RABV cannot spread beyond initially infected cells (Takayama-Ito et al., 2006; Wickersham et al., 2007; Beier et al., 2013). But as for HEP-Flury, G protein could induce the apoptosis of NA cells (Liu et al., 2014; Yang et al., 2015b), consequently limiting the viral spread in CNS (Lay et al., 2003; Sarmento et al., 2005, 2006). Here the results indicated that the viral replication mainly affected efficiency of cell-to-cell spread but not apoptosis-inducing ability.

Analysis of viral RNAs and proteins showed that the ratio of N, P, M, or G mRNA in one transcription was tightly related with gene position though their absolute values were mainly affected by viral replication. We speculated that gene position mainly regulated the gene ratio in one transcription. Moreover, we found that the ratios of leader RNA and L mRNA increased when P gene was translocated. Leader RNA can bind La protein which may inhibit cellular RNA synthesis (Kurilla et al., 1984), thus inhibited the viral replication. And L protein of an attenuated vaccine strain SAD B19 can bind to a dynein light chain 1 (DLC1) acted as a transcription enhancer (Bauer et al., 2015). Both of them decreased the viral replication and stimulated the viral transcription.

In summary, these results revealed that P gene rearrangement of RABV was viable. Viral replication was affected by the absolute value of N protein which was regulated by P protein. Viral lethality in suckling mice was consistent with the ratio of P mRNA in one complete transcription. The protection of mice induced by viruses was related to the antibody titer 5 weeks post-inoculation which might be regulated by G protein. However, the ability to induce cell apoptosis and viral spread were not only related to the viral replication but also to the ratio of related gene which was consistent with the gene position. Subsequently, based on these, we could construct the optimizing RABV as a vaccine candidate and RABV was found lack the mechanism for homologous recombination, this should be an irreversible and stably approach.

AUTHOR CONTRIBUTIONS

MM performed the research and wrote the article; TL, QT, YW, and JL performed the technique of molecular biology; JZ and QZ did the animal experiments; JP provided analysis tools; YL contributed reagents/materials, and XG designed the research and assisted correction of the article.

FUNDING

This study was partially supported by the National Program on Key Research Project of China (No.2016YFD0500400), National Nature Science Foundation of China (No.31172322), Nature Science Foundation of Guangdong (No.2015A03031103), and Special Fund for Agro-Scientific Research in the Public Interest (No.201103032).

ACKNOWLEDGMENTS

We thank the members of our laboratory for helpful discussions and critical comments on the manuscript. Thanks to the HAIDA GROUP for the support of instruments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.00120/full#supplementary-material

FIGURE S1 | Gene expression in NA cells. NA cells were infected with rHEP-Flury, P1, P3 or P4 at a MOI of 3 for 24 h **(A)** and 48 h **(B)** at 34°C. Then RNAs level and viral structural protein were analyzed by qRT-PCR and western blotting respectively. The relative amount of individual viral RNAs was normalized in relation to the housekeeping gene β-actin. Individual RNA ratio were calculated relative to all structural genes plus leader RNA: leader RNA+N+P+M+G+L. Viral structural protein were quantified by Western blotting with monoclonal antibody against RABV N, P, M, G and a monoclonal antibody against actin. Densitometry of the western blotting was analyzed with the Image-Pro Plus 6.0 software. Data are mean ± SD. $n = 3$.

REFERENCES

Albertini, A. A., Ruigrok, R. W., and Blondel, D. (2011). Rabies virus transcription and replication. *Adv. Virus Res.* 79, 1–22. doi: 10.1016/B978-0-12-387040-7.00001-9

Ball, L. A., Pringle, C. R., Flanagan, B., Perepelitsa, V. P., and Wertz, G. W. (1999). Phenotypic consequences of rearranging the P, M, and G genes of vesicular stomatitis virus. *J. Virol.* 73, 4705–4712.

Bauer, A., Nolden, T., Nemitz, S., Perlson, E., and Finke, S. (2015). A dynein light chain 1 binding motif in rabies virus polymerase l protein plays a role in microtubule reorganization and viral primary transcription. *J. Virol.* 89, 9591–9600. doi: 10.1128/JVI.01298-15

Beier, K. T., Saunders, A. B., Oldenburg, I. A., Sabatini, B. L., and Cepko, C. L. (2013). Vesicular stomatitis virus with the rabies virus glycoprotein directs retrograde transsynaptic transport among neurons in vivo. *Front. Neural Circuits* 7:11. doi: 10.3389/fncir.2013.00011

Blondel, D., Regad, T., Poisson, N., Pavie, B., Harper, F., Pandolfi, P. P., et al. (2002). Rabies virus P and small P products interact directly with PML and reorganize PML nuclear bodies. *Oncogene* 21, 7957–7970. doi: 10.1038/sj.onc.1205931

Brzozka, K., Finke, S., and Conzelmann, K. K. (2005). Identification of the rabies virus alpha/beta interferon antagonist: phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *J. Virol.* 79, 7673–7681. doi: 10.1128/JVI.79.12.7673-7681.2005

Care, N. R. C. U., and Animals, A. U. O. L. (2011). *Guide for the Care and Use of Laboratory Animals*. Washington, DC: National Academies Press.

Choi, J., Yang, D. K., Kim, H. H., Jo, H. Y., Choi, S. S., Kim, J. T., et al. (2015). Application of recombinant adenoviruses expressing glycoprotein or nucleoprotein of rabies virus to Korean raccoon dogs. *Clin. Exp. Vaccine Res.* 4, 189–194. doi: 10.7774/cevr.2015.4.2.189

Chopy, D., Detje, C. N., Lafage, M., Kalinke, U., and Lafon, M. (2011). The type I interferon response bridles rabies virus infection and reduces pathogenicity. *J. Neurovirol.* 17, 353–367. doi: 10.1007/s13365-011-0041-6

Davis, B. M., Rall, G. F., and Schnell, M. J. (2015). Everything you always wanted to know about rabies virus (but were afraid to ask). *Annu. Rev. Virol.* 2, 451–471. doi: 10.1146/annurev-virology-100114-055157

Emerson, S. U., and Yu, Y. (1975). Both NS and L proteins are required for in vitro RNA synthesis by vesicular stomatitis virus. *J. Virol.* 15, 1348–1356.

Faber, M., Faber, M. L., Papaneri, A., Bette, M., Weihe, E., Dietzschold, B., et al. (2005). A single amino acid change in rabies virus glycoprotein increases virus spread and enhances virus pathogenicity. *J. Virol.* 79, 14141–14148. doi: 10.1128/JVI.79.22.14141-14148.2005

Faber, M., Pulmanausahakul, R., Hodawadekar, S. S., Spitsin, S., McGettigan, J. P., Schnell, M. J., et al. (2002). Overexpression of the rabies virus glycoprotein results in enhancement of apoptosis and antiviral immune response. *J. Virol.* 76, 3374–3381. doi: 10.1128/JVI.76.7.3374-3381.2002

Finke, S., and Conzelmann, K. K. (2003). Dissociation of rabies virus matrix protein functions in regulation of viral RNA synthesis and virus assembly. *J. Virol.* 77, 12074–12082. doi: 10.1128/JVI.77.22.12074-12082.2003

Finke, S., and Conzelmann, K. K. (2005). Replication strategies of rabies virus. *Virus Res.* 111, 120–131. doi: 10.1016/j.virusres.2005.04.004

Finke, S., Mueller-Waldeck, R., and Conzelmann, K. K. (2003). Rabies virus matrix protein regulates the balance of virus transcription and replication. *J. Gen. Virol.* 84, 1613–1621. doi: 10.1099/vir.0.19128-0

Flanagan, E. B., Ball, L. A., and Wertz, G. W. (2000). Moving the glycoprotein gene of vesicular stomatitis virus to promoter-proximal positions accelerates and enhances the protective immune response. *J. Virol.* 74, 7895–7902. doi: 10.1128/JVI.74.17.7895-7902.2000

Fouquet, B., Nikolic, J., Larrous, F., Bourhy, H., Wirblich, C., Lagaudriere-Gesbert, C., et al. (2015). Focal adhesion kinase is involved in rabies virus infection through its interaction with viral phosphoprotein P. *J. Virol.* 89, 1640–1651. doi: 10.1128/JVI.02602-14

Ghanem, A., Kern, A., and Conzelmann, K. K. (2012). Significantly improved rescue of rabies virus from cDNA plasmids. *Eur. J. Cell Biol.* 91, 10–16. doi: 10.1016/j.ejcb.2011.01.008

Inoue, K., Shoji, Y., Kurane, I., Iijima, T., Sakai, T., and Morimoto, K. (2003). An improved method for recovering rabies virus from cloned cDNA. *J. Virol. Methods* 107, 229–236. doi: 10.1016/S0166-0934(02)00249-5

Kammouni, W., Wood, H., Saleh, A., Appolinario, C. M., Fernyhough, P., and Jackson, A. C. (2015). Rabies virus phosphoprotein interacts with mitochondrial Complex I and induces mitochondrial dysfunction and oxidative stress. *J. Neurovirol.* 21, 370–382. doi: 10.1007/s13365-015-0320-8

Kassis, R., Larrous, F., Estaquier, J., and Bourhy, H. (2004). Lyssavirus matrix protein induces apoptosis by a TRAIL-dependent mechanism involving caspase-8 activation. *J. Virol.* 78, 6543–6555. doi: 10.1128/JVI.78.12.6543-6555.2004

Kuang, Y., Lackay, S. N., Zhao, L., and Fu, Z. F. (2009). Role of chemokines in the enhancement of BBB permeability and inflammatory infiltration after rabies virus infection. *Virus Res.* 144, 18–26. doi: 10.1016/j.virusres.2009.03.014

Kurilla, M. G., Cabradilla, C. D., Holloway, B. P., and Keene, J. D. (1984). Nucleotide sequence and host La protein interactions of rabies virus leader RNA. *J. Virol.* 50, 773–778.

Lay, S., Préhaud, C., Dietzschold, B., and Lafon, M. (2003). Glycoprotein of nonpathogenic rabies viruses is a major inducer of apoptosis in human jurkat T cells. *Ann. N. Y. Acad. Sci.* 1010, 577–581. doi: 10.1196/annals.1299.108

Li, J., Faber, M., Papaneri, A., Faber, M. L., McGettigan, J. P., Schnell, M. J., et al. (2006). A single immunization with a recombinant canine adenovirus expressing the rabies virus G protein confers protective immunity against rabies in mice. *Virology* 356, 147–154. doi: 10.1016/j.virol.2006.07.037

Liu, P., Yang, J., Wu, X., and Fu, Z. F. (2004). Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. *J. Gen. Virol.* 85, 3725–3734. doi: 10.1099/vir.0.80325-0

Liu, X., Yang, Y., Sun, Z., Chen, J., Ai, J., Dun, C., et al. (2014). A recombinant rabies virus encoding two copies of the glycoprotein gene confers protection in dogs against a virulent challenge. *PLoS ONE* 9:e87105. doi: 10.1371/journal.pone.0087105

Luo, J., Zhao, J., Tian, Q., Mo, W., Wang, Y., Chen, H., et al. (2016). A recombinant rabies virus carrying GFP between N and P affects viral transcription in vitro. *Virus Genes* 52, 379–387. doi: 10.1007/s11262-016-1313-2

Marschalek, A., Finke, S., Schwemmle, M., Mayer, D., Heimrich, B., Stitz, L., et al. (2009). Attenuation of rabies virus replication and virulence by picornavirus internal ribosome entry site elements. *J. Virol.* 83, 1911–1919. doi: 10.1128/JVI.02055-08

Morimoto, K., Kawai, A., and Mifune, K. (1992). Comparison of rabies virus G proteins produced by cDNA-transfected animal cells that display either inducible or constitutive expression of the gene. *J. Gen. Virol.* 73(Pt 2), 335–345. doi: 10.1099/0022-1317-73-2-335

Morimoto, K., Kawai, A., Sato, Y., and Ohkubo, A. (2011). A unique transcription mode of rabies virus high egg passage-Flury strain detected in infected baby hamster kidney-21 cells. *Microbiol. Immunol.* 55, 558–564. doi: 10.1111/j.1348-0421.2011.00354.x

Morimoto, K., Shoji, Y., and Inoue, S. (2005). Characterization of P gene-deficient rabies virus: propagation, pathogenicity and antigenicity. *Virus Res.* 111, 61–67. doi: 10.1016/j.virusres.2005.03.011

Niu, X., Tang, L., Tseggai, T., Guo, Y., and Fu, Z. F. (2013). Wild-type rabies virus phosphoprotein is associated with viral sensitivity to type I interferon treatment. *Arch. Virol.* 158, 2297–2305. doi: 10.1007/s00705-013-1743-2

Niu, X., Wang, H., and Fu, Z. F. (2011). Role of chemokines in rabies pathogenesis and protection. *Adv. Virus Res.* 79, 73–89. doi: 10.1016/B978-0-12-387040-7.00005-6

Okumura, A., and Harty, R. N. (2011). Rabies virus assembly and budding. *Adv. Virus Res.* 79, 23–32. doi: 10.1016/B978-0-12-387040-7.00002-0

Pattnaik, A. K., and Wertz, G. W. (1990). Replication and amplification of defective interfering particle RNAs of vesicular stomatitis virus in cells expressing viral proteins from vectors containing cloned cDNAs. *J. Virol.* 64, 2948–2957.

Peng, J., Zhu, S., Hu, L., Ye, P., Wang, Y., Tian, Q., et al. (2016). Wild-type rabies virus induces autophagy in human and mouse neuroblastoma cell lines. *Autophagy* 12, 1704–1720. doi: 10.1080/15548627.2016.1196315

Préhaud, C., Lay, S., Dietzschold, B., and Lafon, M. (2003). Glycoprotein of nonpathogenic rabies viruses is a key determinant of human cell apoptosis. *J. Virol.* 77, 10537–10547. doi: 10.1128/JVI.77.19.10537-10547.2003

Pringle, C. R. (1982). The genetics of vesiculoviruses. *Arch. Virol.* 72, 1–34. doi: 10.1007/BF01314447

Pringle, C. R. (1996). Virus taxonomy 1996 - a bulletin from the Xth International Congress of Virology in Jerusalem. *Arch. Virol.* 141, 2251–2256. doi: 10.1007/BF01718231

Rieder, M., Brzozka, K., Pfaller, C. K., Cox, J. H., Stitz, L., and Conzelmann, K. K. (2011). Genetic dissection of interferon-antagonistic functions of rabies virus phosphoprotein: inhibition of interferon regulatory factor 3 activation is important for pathogenicity. *J. Virol.* 85, 842–852. doi: 10.1128/JVI.01427-10

Rieder, M., and Conzelmann, K. K. (2011). Interferon in rabies virus infection. *Adv. Virus Res.* 79, 91–114. doi: 10.1016/B978-0-12-387040-7.00006-8

Sarmento, L., Li, X.-Q., Howerth, E., Jackson, A. C., and Fu, Z. F. (2005). Glycoprotein-mediated induction of apoptosis limits the spread of attenuated rabies viruses in the central nervous system of mice. *J. Neurovirol.* 11, 571–581. doi: 10.1080/13550280500385310

Sarmento, L., Tseggai, T., Dhingra, V., and Fu, Z. F. (2006). Rabies virus-induced apoptosis involves caspase-dependent and caspase-independent pathways. *Virus Res.* 121, 144–151. doi: 10.1016/j.virusres.2006.05.002

Shoji, Y., Inoue, S., Nakamichi, K., Kurane, I., Sakai, T., and Morimoto, K. (2004). Generation and characterization of P gene-deficient rabies virus. *Virology* 318, 295–305. doi: 10.1016/j.virol.2003.10.001

Takayama-Ito, M., Inoue, K., Shoji, Y., Inoue, S., Iijima, T., Sakai, T., et al. (2006). A highly attenuated rabies virus HEP-Flury strain reverts to virulent by single amino acid substitution to arginine at position 333 in glycoprotein. *Virus Res.* 119, 208–215. doi: 10.1016/j.virusres.2006.01.014

Thoulouze, M.-I., Lafage, M., Yuste, V. J., Baloul, L., Edelman, L., Kroemer, G., et al. (2003). High level of Bcl-2 counteracts apoptosis mediated by a live rabies virus vaccine strain and induces long-term infection. *Virology* 314, 549–561. doi: 10.1016/S0042-6822(03)00491-4

Ugolini, G. (2011). Rabies virus as a transneuronal tracer of neuronal connections. *Adv. Virus Res.* 79, 165–202. doi: 10.1016/B978-0-12-387040-7.00010-X

Vidy, A., Chelbi-Alix, M., and Blondel, D. (2005). Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. *J. Virol.* 79, 14411–14420. doi: 10.1128/JVI.79.22.14411-14420.2005

Wang, Y., Tian, Q., Xu, X., Yang, X., Luo, J., Mo, W., et al. (2014). Recombinant rabies virus expressing IFNalpha1 enhanced immune responses resulting in its attenuation and stronger immunogenicity. *Virology* 46, 621–630. doi: 10.1016/j.virol.2014.09.010

Wertz, G. W., Perepelitsa, V. P., and Ball, L. A. (1998). Gene rearrangement attenuates expression and lethality of a nonsegmented negative strand RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3501–3506. doi: 10.1073/pnas.95.7.3501

Wickersham, I. R., Finke, S., Conzelmann, K. K., and Callaway, E. M. (2007). Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* 4, 47–49. doi: 10.1038/nmeth999

Wirblich, C., and Schnell, M. J. (2011). Rabies virus (RV) glycoprotein expression levels are not critical for pathogenicity of RV. *J. Virol.* 85, 697–704. doi: 10.1128/JVI.01309-10

Wu, X., and Rupprecht, C. E. (2008). Glycoprotein gene relocation in rabies virus. *Virus Res.* 131, 95–99. doi: 10.1016/j.virusres.2007.07.018

Yamaoka, S., Ito, N., Ohka, S., Kaneda, S., Nakamura, H., Agari, T., et al. (2013). Involvement of the rabies virus phosphoprotein gene in neuroinvasiveness. *J. Virol.* 87, 12327–12338. doi: 10.1128/JVI.02132-13

Yang, X. F., Peng, J. J., Liang, H. R., Yang, Y. T., Wang, Y. F., Wu, X. W., et al. (2014). Gene order rearrangement of the M gene in the rabies virus leads to slower replication. *Virusdisease* 25, 365–371. doi: 10.1007/s13337-014-0220-1

Yang, Y., Huang, Y., Gnanadurai, C. W., Cao, S., Liu, X., Cui, M., et al. (2015a). The inability of wild-type rabies virus to activate dendritic cells is dependent on the glycoprotein and correlates with its low level of the de novo-synthesized leader RNA. *J. Virol.* 89, 2157–2169. doi: 10.1128/JVI.02092-14

Yang, Y., Liu, W., Yan, G., Luo, Y., Zhao, J., Yang, X., et al. (2015b). iTRAQ protein profile analysis of neuroblastoma (NA) cells infected with the rabies viruses rHep-Flury and Hep-dG. *Front. Microbiol.* 6:691. doi: 10.3389/fmicb.2015.00691

Zhao, L., Toriumi, H., Kuang, Y., Chen, H., and Fu, Z. F. (2009). The roles of chemokines in rabies virus infection: overexpression may not always be beneficial. *J. Virol.* 83, 11808–11818. doi: 10.1128/JVI.01346-09

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Mei, Long, Zhang, Zhao, Tian, Peng, Luo, Wang, Lin and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 June 2017
doi: 10.3389/fmicb.2017.01091

Viral Population Changes during Murine Norovirus Propagation in RAW 264.7 Cells

Takuya Kitamoto[1†], Reiko Takai-Todaka[2†], Akiko Kato[1], Kumiko Kanamori[3], Hirotaka Takagi[4], Kazuhiro Yoshida[3], Kazuhiko Katayama[2,5] and Akira Nakanishi[1,3]**

[1] Laboratory of Radiation Safety, National Center for Geriatrics and Gerontology, Obu, Japan, [2] Laboratory of Gastroenteritis Viruses, Virology II, National Institute for Infectious Diseases, Musashimurayama, Japan, [3] Section of Gene Therapy, Department of Aging Intervention, National Center for Geriatrics and Gerontology, Obu, Japan, [4] Division of Biosafety Control and Research, National Institute for Infectious Diseases, Tokyo, Japan, [5] Laboratory of Viral Infection I, Graduate School of Infection Control Sciences, Kitasato Institute for Life Sciences, Kitasato University, Tokyo, Japan

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Yashpal S. Malik,
Indian Veterinary Research Institute, India
Joana Rocha-Pereira,
KU Leuven, Belgium
Abimbola O. Kolawole,
University of Michigan Health System, United States

***Correspondence:**
Akira Nakanishi
nakanish@ncgg.go.jp
Kazuhiko Katayama
katayama@lisci.kitasato-u.ac.jp

[†] These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *08 January 2017*
Accepted: *30 May 2017*
Published: *15 June 2017*

Citation:
Kitamoto T, Takai-Todaka R, Kato A, Kanamori K, Takagi H, Yoshida K, Katayama K and Nakanishi A (2017) Viral Population Changes during Murine Norovirus Propagation in RAW 264.7 Cells. Front. Microbiol. 8:1091. doi: 10.3389/fmicb.2017.01091

Laboratory adaptation of viruses is an essential technique for basic virology research, including the generation of attenuated vaccine strains, although the principles of cell adaptation remain largely unknown. Deep sequencing of murine norovirus (MuNoV) S7 during serial passages in RAW264.7 cells showed that the frequencies of viral variants were altered more dynamically than previously reported. Serial passages of the virus following two different multiplicity of infections gave rise to distinct haplotypes, implying that multiple cell-adaptable sequences were present in the founder population. Nucleotide variants lost during passage were assembled into a viral genome representative of that prior to cell adaptation, which was unable to generate viral particles upon infection in cultured cells. In addition, presence of the reconstructed genome interfered with production of infectious particles from viruses that were fully adapted to *in vitro* culture. Although the key nucleotide changes dictating cell adaptation of MuNoV S7 viral infection are yet to be elucidated, our results revealed the elaborate interplay among selected sequences of viral variants better adapted to propagation in cell culture. Such knowledge will be instrumental in understanding the processes necessary for the laboratory adaptation of viruses, especially to those without relevant cell culture systems.

Keywords: norovirus, RAW264.7 cells, cell adaptation, mouse, *Calicivirus*

INTRODUCTION

Virus propagation *in vitro* using cultured cells is essential for virology research and for the production of attenuated virus for use in vaccines, including those for measles, polio, and rabies (Minor, 2015); however, the processes shaping viral adaptation to cell culture conditions remain poorly understood.

Norovirus (NoV) is the most prevalent cause of viral gastroenteritis worldwide. The virus belongs to *Caliciviridae* family and consists of a single-stranded, positive-sense, ∼7.6-kb RNA genome that includes a 3′ poly(A) tail. The genome harbors three open reading frames (ORFs)—ORF1 encodes a non-structural (NS) polyprotein that constitutes NS1-2, NS3, NS4, NS5, NS6, and NS7 proteins, whereas ORF2 and ORF3 encode the major and minor capsid proteins VP1

and VP2, respectively. In murine NoV (MuNoV), the accessory viral protein VF1 is encoded by the ORF2 region via an alternative VP1 reading frame (McFadden et al., 2011).

Due to the lack of a permissive cell line, it is currently infeasible to propagate human NoV (HuNoV) using conventional *in vitro* culture systems (Duizer et al., 2004). Limited growth of HuNoV in immunodeficient mice and in human B-cell lines co-cultured with intestinal bacteria has been reported (Taube et al., 2013; Jones et al., 2014). Recent advances in reverse-genetics (Katayama et al., 2014) and development of an enteroid culture system using intestinal stem cells have proven effective for HuNoV propagation *in vitro* (Ettayebi et al., 2016); however, this culture system is impractical for general use owing to its high cost. As such, MuNoV is often used as a surrogate model virus to study NoV because of its ability to proliferate in RAW264.7 cell cultures, as well as murine dendritic cells (Wobus et al., 2004), and of its similarity to HuNoV being infectious agents to gastrointestinal tract. The prototypic MuNoV strain MNV-1 was originally discovered as a lethal agent in *RAG2/STAT1*$^{-/-}$ mice (Karst et al., 2003), although MNV-1 infections in wild-type mice caused little pathological changes and were quickly cleared by the host immune system. In comparison, a more prevalent MuNoV stain, CR3, causes persistent infection in the gastrointestinal tract of laboratory mice and is mostly benign to the host, even in immunodeficient mice (Thackray et al., 2007). The MuNoV S7 strain, which is the closest relative to CR3, also appeared to induce no pathological effects (Kitajima et al., 2009); however, whether the strain can cause persistent infection in mice is not known.

Cell adaptation of MNV-1 is associated with attenuated viral pathogenicity in host animals (Wobus et al., 2004). For instance, the V11I and E296K mutations in NS4 and VP1, respectively, are associated with an inability to cause lethal infection in immunodeficient mice, although the potential for growth in RAW264.7 cells appeared unaltered or slightly enhanced (Wobus et al., 2004; Bailey et al., 2008). MNV-1 genomic alterations during cell passages have been well-documented by Sanger sequencing (Wobus et al., 2004; Bailey et al., 2008) and partially by deep sequencing (Mauroy et al., 2016); however, a detailed analysis of these changes in other MuNoV strains and their effect on cell adaptation has yet to be determined.

Here, we report on the detailed examination of MuNoV S7 population changes during cell passage using deep sequencing at two different multiplicity of infections (MOIs). The frequencies of sequence variations were monitored at each passage and linkages analyzed by haplotype reconstruction. Additionally, nucleotide variants lost during cell passage were assembled and used to construct a single genome, which was then examined for its ability to generate infectious particles and possible interactions with genomic sequences fully adapted to *in vitro* propagation. These results revealed dynamic associations between viral population changes in response to cell culture conditions and the complex interplay between viral variants during the selection of sequences better suited for *in vitro* propagation.

MATERIALS AND METHODS

Cells and Viruses

RAW264.7 cells (ATCC TIB-71; American Type Culture Collection, Manassas, VA, United States) were cultured according to the manufacturer's instructions. The MuNoV S7 strain was kindly provided by Dr. Yukinobu Tohya (Nihon University, Tokyo, Japan) and grown in RAW264.7 cells. Viral preparations derived from cell cultures passaged two or three times were designated as P2 and P3 virus, respectively.

MuNoV infection titers were examined according to the 50% cell culture infective dose (CCID50) as previously described (Katayama et al., 2014). Viral RNA copy number was quantified by real-time reverse transcription polymerase chain reaction (RT-PCR) to approximate viral particle number. Briefly, viral RNA was extracted using QIAamp viral RNA mini kit (Qiagen, Hilden, Germany) and amplified with iScript One-Step RT-PCR kit using SYBR Green (Bio-Rad, Hercules, CA, United States) and the primers, MNV 6082-6108 FW and MNV 6272-6246 RV (Supplemental Table 1), with CFX96 Real-time PCR detection system (Bio-Rad). Fixed amounts of *in vitro*-transcribed MuNoV RNA served as a standard (see "Generation of Recombinant MuNoV").

The number of viral particles required to achieve one infectious event was assessed using the P13 virus prepared from RAW264.7 cells passaged 10 times after infection with P3 virus. Particle numbers and infectious titers were estimated by quantifying viral RNA copy number and the CCID50, respectively. In our preliminary assessments, ∼500 viral RNA copies were equivalent to one CCID50 unit, although others have shown that the ratio of infectious particles to viral genomes was approximately 1:100 (Baert et al., 2008; Fischer et al., 2015). This five-fold difference could result from differences in quantification method, strain characteristics, or cell culture conditions.

Sequence Determination of Viral RNA by Sanger Sequencing

Extracted RNA from P2 and P3 MuNoV-containing culture supernatants was resuspended in AVE buffer (Qiagen) and used to generate cDNA in triplicate reactions with either SuperScript III reverse transcriptase (Thermo Fisher Scientific, Waltham, MA, United States) or ReverTra Ace (Toyobo, Osaka, Japan). The resulting cDNA was amplified in triplicate reactions using PrimeStar enzymes (Takara, Otsu, Japan) and five primer sets (Supplemental Table 2), each generating fragments that partially overlapped with the adjacent fragments; thus, nine samples were generated each genomic fragment. The amplified fragments were then separated by gel electrophoresis, isolated, and sequenced with the Big Dye ready reaction kit 3.1 (Thermo Fisher Scientific) and a 3130xl Genetic Analyzer (Thermo Fisher Scientific).

Preparation of Viral RNA for Next-Generation Sequencing (NGS)

Two different viral populations were generated by passaging the P2 virus in RAW264.7 cells grown in a 6-cm dish. The first was applied to cells at an MOI $>$ 5 and incubated for

∼24 h, at which point many of the cells had died, although more extensive cell death was observed at later passages. Cell supernatant (∼4 mL) was cleared of cell debris by two successive 10 min centrifugations—at 700 × *g* and then 12,000 × *g*—at 4°C. The virus was then concentrated by ultracentrifugation with a SW50.1 rotor (Beckman Coulter, Fullerton, CA, United States) at 45,000 rpm for 2 h at 4°C, and the resulting pellet was resuspended in 0.1 mL AVE buffer (Qiagen) for RNA extraction with the QIAamp viral RNA mini kit (Qiagen).

The other viral population was generated by infecting the cells at an MOI of 1 CCID50 (∼500 viral particles)/cell and the supernatant was collected at 48-h post infection (hpi). Cell supernatant (∼4 mL) from each passage was concentrated as described above, the RNA extracted using Isogen II (Takara), and then subsequently resuspended in RNase-free water.

Reverse Transcription, Double-Stranded DNA Synthesis, Library Preparation, and Deep Sequencing

Viral RNA was used to synthesize cDNA with Superscript III reverse transcriptase (Thermo Fisher Scientific), which was then amplified using PrimeStar GXL enzymes (Takara) and two sets of primers, MNV-1S 14-32/MNV-7A 5361-5380 and MNV-4S 4989-5008/TX30SXN (Supplemental Table 1). The resulting fragments encompassed the proximal and distal halves of the viral genome, respectively, with an overlap of >300 nucleotides. The DNA fragments were gel-isolated and 1 ng used to prepare a cDNA library with the Nextera XT DNA Sample Prep Kit (Illumina, San Diego, CA, United States) according to manufacturer's instructions. Briefly, DNA was fragmented and tagged by the Nextera XT transposome and then used as a template in a 50-μL, 12-cycle PCR. The amplified DNA was processed as outlined in the Nextera XT protocol and purified with AMPureXP beads (Beckman).

The quality of the purified DNA libraries was assessed on a MultiNA MCE-202 Bioanalyzer (Shimadzu Corporation, Kyoto, Japan). Nucleotide sequencing was performed on an Illumina MiSeq sequencer with a MiSeq Reagent Kit v2 (Illumina) to generate 151 paired-end reads.

Data Processing and Analysis

Short-reads were trimmed and mapped to the MuNoV reference sequence (GenBank: AB435515.1) using CLC Genomics Workbench 4.65 (CLC Bio, Cambridge, MA, United States) with default alignment settings. BAM files were exported from the CLC Genomics Workbench and analyzed using SAMtools v1.3.1 (Li et al., 2009) to extract sequence coverage and relevant statistics (Supplemental Presentation 1). Major nucleotide variants were called using CLC Genomics Workbench and their frequencies sampled and extracted using SAMtools (Supplemental Table 4). QuasiRecomb version 1.2 (Topfer et al., 2013) was used to reconstruct the viral haplotypes from sequencing data in BAM files. Local haplotype reconstruction across genes was performed using the default setting and conservative parameters to determine the inferred haplotypes and estimate their frequencies.

Construction of DNA

Unless noted, PrimeStar enzymes (Takara) and the InFusion system (Takara) were used for PCR-fragment generation and fragment insertion into plasmid DNA, respectively. Primers used in this study are listed in Supplemental Table 1. All construct sequences were confirmed by Sanger sequencing.

All genomic variations different from the MuNoV S7 cDNA reference sequence (GenBank: AB435515.1) or PP3 detected by deep sequencing (Supplemental Table 4) were assembled and synthesized as a single artificial genome—termed the PP2 composite sequence (PP2com). The nucleotide differences between PP2com and PP3 spanned from nucleotides 87 to 6752 in the viral genome. The EcoRI and BstBI sites positioned at –73 and 6996 bp in pMNV S7, respectively, (Katayama et al., 2014) were used to replace the original MuNoV sequence with that of PP2com. The 5′-proximal sequence up to the EcoRI site was also included in the synthesis.

pMNV ORF1 PP2, harboring the ORF1 from PP2com, was constructed by exchanging the PP3 sequence between the XhoI and BstBI restriction sites with the ORF2 and ORF3 from pMNV PP2, as well as a small region of the RNA-dependent RNA polymerase-coding sequence that included silent mutations at +4745 and +4877 bp in the genome.

Similarly, pMNV ORF23 PP2, harboring the ORF2 and ORF3 from PP2com, was generated by replacing the pMNV S7 fragment between the XhoI and BstBI restriction sites in with the PP2 sequence from pMNV PP2.

pT7 MNV S7 was generated from pMNV S7 (Katayama et al., 2014) by replacing the EF1-alpha promoter between the SspI and MluI restriction sites and with the T7-promoter sequence. First, two overlapping PCR fragments, A and B, were generated by PCR using pMNV S7 as the template with the primers, SspI FW and T7 MNV RV for A, and MluI RV and T7 MNV FW for B (Supplemental Table 1). The two fragments were then mixed and used as the templates for making insert by PCR using SspI FW and MluI RV as the primers for replacing SspI-MluI fragment of pMNV S7. The resultant pT7 MNVS7 harbors a truncated T7 promoter sequence (5′-TAATACGACTCACTATA-3′) placed proximal to the MNV genomic cDNA sequence for generate RNA with 5′ ends identical to that of natural MuNoV upon *in vitro* transcription.

Generation of Recombinant MuNoV

Recombinant MuNoV was produced using a plasmid-based reverse-genetics system as described previously (Katayama et al., 2014). The pMNV S7, pMNV PP2, pMNV ORF1 PP2, and pMNV ORF23 PP2 plasmids were transfected individually or in combination into 293T cells cultured in 35-mm dishes by mixing 4 μg DNA, 8 μL P3000 reagent, and 12 μL Lipofectamine 3000 (Thermo Fisher Scientific) in 400 μL Opti-MEM (Thermo Fisher Scientific). Supernatants from transfected cultures were collected 48-h post-transfection and used to infect RAW264.7 cells. The transfected cells were also collected to confirm viral protein expression by Western blotting.

In addition, recombinant MuNoV was also generated using a RNA-based reverse-genetics system as previously described (Arias et al., 2012). Briefly, a series of pT7 MNV constructs was linearized with AscI and used as the template for *in vitro* transcription with the T7 RiboMax express large-scale RNA-production system (Promega, Fitchburg, WI, United States). The synthesized RNA was purified using a Megaclear kit (Thermo Fisher Scientific) followed by 5′-end capping by ScriptCap (Illumina). The capped RNA was then re-purified and used to transfect 293T cells in 35-mm dishes with 4 μg RNA and 8 μL Lipofectamine 2000 (Thermo Fisher Scientific). The viral supernatant was collected 24 h later and used to infect RAW 264.7 cells. The virus was passaged once in RAW264.7 cells to obtain a sufficient titer for cell-infection experiments.

Western Blotting

Viral proteins from transfected 293T cells were analyzed by Western blotting as previously described (Haga et al., 2016). Briefly, the cell pellet was resuspended in calcium- and magnesium-free Dulbecco's Phosphate-Buffered Saline (PBS^-), disrupted by sonication, and centrifuged to remove cell debris. The protein in each sample was quantified and then denatured with SDS-dye buffer (Wako Chemicals, Tokyo, Japan). Aliquots containing about ∼10 μg protein were loaded into each lane, separated by 5–20% SDS-PAGE, and transferred onto a polyvinylidene fluoride (PVDF) membrane using a Trans-Blot system (Bio-Rad). The membrane was blocked with solution containing 2.5% skim milk and 0.5× PVDF blocking buffer (Toyobo), followed by rabbit anti-NS1-2 in solution 1 from the Can Get Signal Kit (Toyobo). After washing, immunoreactive bands were detected with horseradish peroxidase (HRP)-conjugated anti-rabbit IgG antibody in solution 2. Chemiluminescence was detected by ImmunoStar (Wako Chemicals) and recorded using a LAS4000 (GE Healthcare, Little Chalfont, United Kingdom). The membrane was then stripped with Restore Plus Western blot stripping buffer (Thermo Fisher Scientific) and reprobed with HRP-conjugated guinea pig anti-NS7 and mouse anti-actin (Wako Chemicals).

Immunofluorescent Staining

Supernatants (∼0.5 mL) from cells transfected with MuNoV genome constructs were used to infect RAW264.7 cells grown on coverslips in each well of a 12-well plate. The cells were fixed 48 h later with 4% paraformaldehyde (pH 7.5) and examined for viral protein expression by immunocytochemistry. Mixtures of primary anti-sera, guinea pig anti-VPg, and rabbit anti-VP1 in PBS^- supplemented with 0.5% Triton X-100 were used to detect the respective proteins, followed by AlexaFluor 488-conjugated anti-guinea pig IgG and AlexaFluor555-conjugated anti-rabbit IgG secondary antibodies. Samples were embedded in Vectashield mounting medium (Vector Labs., Burlingame, CA, United States) containing 4′,6-diamidino-2-phenylindole (DAPI) and imaged with an epifluorescent microscope (BZ9000; Keyence, Osaka, Japan).

RESULTS

The MuNoV S7 PP3 sequence (GenBank: AB435515.1) originated from a molecular clone obtained from the P3 virus that was passaged three times in RAW264.7 cells (**Figure 1**). The recombinant virus generated from the PP3 sequence by the plasmid-based reverse-genetics system grew well in RAW264.7 cells (Katayama et al., 2014); however, Sanger sequencing revealed that the P3 "consensus" sequence differed by more than 55 nucleotides with 11 amino acid-changes from the PP3 sequence (Supplemental Table 3). Similarly the consensus sequence of P2 virus, which was passaged twice in RAW264.7 cells, differed by 48 nucleotides with 11 missense changes from the reference sequence suggesting that both viral preparations contained pools of heterogeneous sequences (Supplemental Table 3). After 10 passages of the P3 virus in RAW264.7 cells (P13 virus), the consensus viral sequence became identical to that of PP3, thereby confirming that this MuNoV S7 sequence was representative of the one adapted to growth in RAW264.7 cells.

Because earlier passages of MuNoV S7 isolates—including the P2 virus–contained variable clones differing from those of PP3, this suggested that the original viral isolate contained multiple variants, which were further refined by continued propagation in RAW264.7 cells. To further delineate the process by which this occurred, viral sequences from each passage were examined by deep sequencing. For this, viral RNA was prepared from the two set of passages inoculated at different MOIs. The first culture was passaged at an MOI > 5 CCID50/cell, assuming that multiple viral clones would likely coexist in a single cell upon inoculation (high-MOI passage). After six passages, viral RNA was extracted from the infected culture supernatants for sequencing and defined as H1 to H6 for passages one to six. The second culture was inoculated at an MOI of 1 CCID50/cell, assuming that ∼63% of cells were infected and 37% of those were infected with only one virus (a low-MOI passage). It should also be noted that MOIs were only valid at inoculation, as much higher MOI would be expected after single round of viral propagation and release at subsequent infection to adjacent cells. Culture supernatants were harvested over nine passages. Samples L1 to L4 represented those collected from passages two to five, whereas L5, L6, and L7 were from passages six, seven, and nine, respectively.

Deep sequencing of the viral RNA samples H1–H6 and L1–L7 revealed dynamic sequence changes in the viral population. High- and low-MOI passages showed a total of 89 and 103 nucleotide differences from the PP3 sequence, respectively, of which 87 were shared. The variations included 17 non-synonymous and 22 missense mutations, including 8 shared among the passages (**Figure 2** and Supplemental Table 4). The proportion of non-PP3 sequences—referred to as PP2 sequences—in each passage is summarized in **Figure 2**, with detailed data provided in Supplemental Table 4. Overall changes in nucleotide frequencies in the low-MOI passage were far more evident than those observed with the high-MOI infection, of which many in low-MOI Passage 9 (sample L7) reached nearly zero, indicating that several variants were lost. In contrast, variant frequencies in

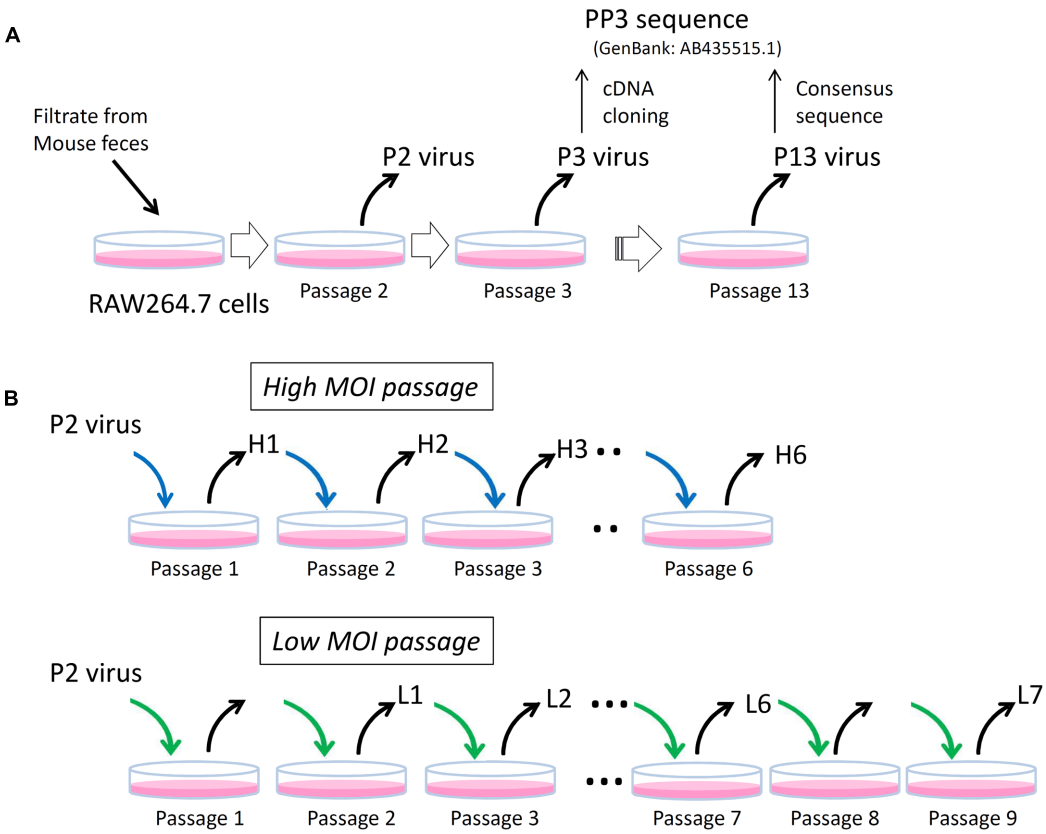


FIGURE 1 | Schematics of experimental layout. **(A)** P2, P3, and P13 viruses. The P2 and P3 viruses are MuNoV preparations passaged in RAW264.7 cells twice (Passage 2) and three times (Passage 3), respectively. Representative molecular clone of MuNoV S7, PP3, was generated by cDNA cloning from P3 virus. After ten passages of P3 virus (Passage 13), which designated as P13 virus, the consensus sequence of the virus became identical to that of PP3. **(B)** High and low MOI passages. Two different MOI conditions were used to inoculate P2 virus and subsequent passages in RAW264.7 cells. The blue and green arrows indicate viral inoculation under MOI > 5CCID50 unit and MOI = 1 CCID50 unit, respectively. Note in "Low MOI passage" number of passages and the samples do not match; no viral samples were taken at passages 1 and 8.

high-MOI samples showed relatively smaller changes. Moreover, further examination of the low-MOI variants revealed temporal shifts in some nucleotide frequencies (Supplemental Table 4).

To investigate nucleotide frequency linkages, we used QuasiRecomb to generate haplotypes (Topfer et al., 2013). QuasiRecomb implements a hidden Markov model to infer viral quasispecies from deep-coverage NGS data using an expectation-maximization algorithm for maximum posterior-parameter estimation and explicitly accounts for paired-end information. Haplotype reconstructions for the full genome, ORF1, ORF2, and ORF3 are shown in **Table 1A**. Computational processing required deep reads and >1000 coverages were prerequisite for distinguishing single-nucleotide polymorphisms and misreads by deep sequencing (Topfer et al., 2013). Not surprisingly, applying this method to samples H1, L3, and L4 with <1000 coverages on average resulted in large numbers of haplotypes, which likely included misreads. Similar findings were observed with H3 and H4, although it was unknown whether this was attributed to the relatively smaller degree of coverage. Results of comparisons from the remaining samples (H2, H5, and H6 from high-MOI passage and L1 and L4–L7 from low-MOI passage) revealed a general trend of fewer identified haplotypes in low-MOI cultures. Moreover, the ORF1 region generated one or two haplotypes during the final passages of both MOIs, whereas the ORF2 region showed 15 and 2 haplotypes for high- and low-MOI cultures, respectively, which could explain the differences in haplotype number observed for the full genomes (29 and 2 for high- and low-MOI passages, respectively).

While the conservative settings only reconstruct major haplotypes and disregard minor ones, the default setting includes minor haplotypes and generates details for changes in haplotype frequencies. However, haplotype reconstructions for ORF1 and the full genome were unsuccessful with the default settings because of inadequate coverage. As such, we only reconstructed the haplotypes with the default setting for ORF2 and ORF3 regions (**Table 1B**). Notably, the ORF2 region harbored a large number of viral haplotypes; 3569 and 5343 in H1 and L1, respectively. In addition, fewer haplotypes were observed in late passages from both MOIs, indicative of haplotype convergence, although more haplotypes were present in H6 as compared to L7. In particular, major haplotype frequencies associated with the ORF2 region observed were 0.691 and 0.059 in H6 and

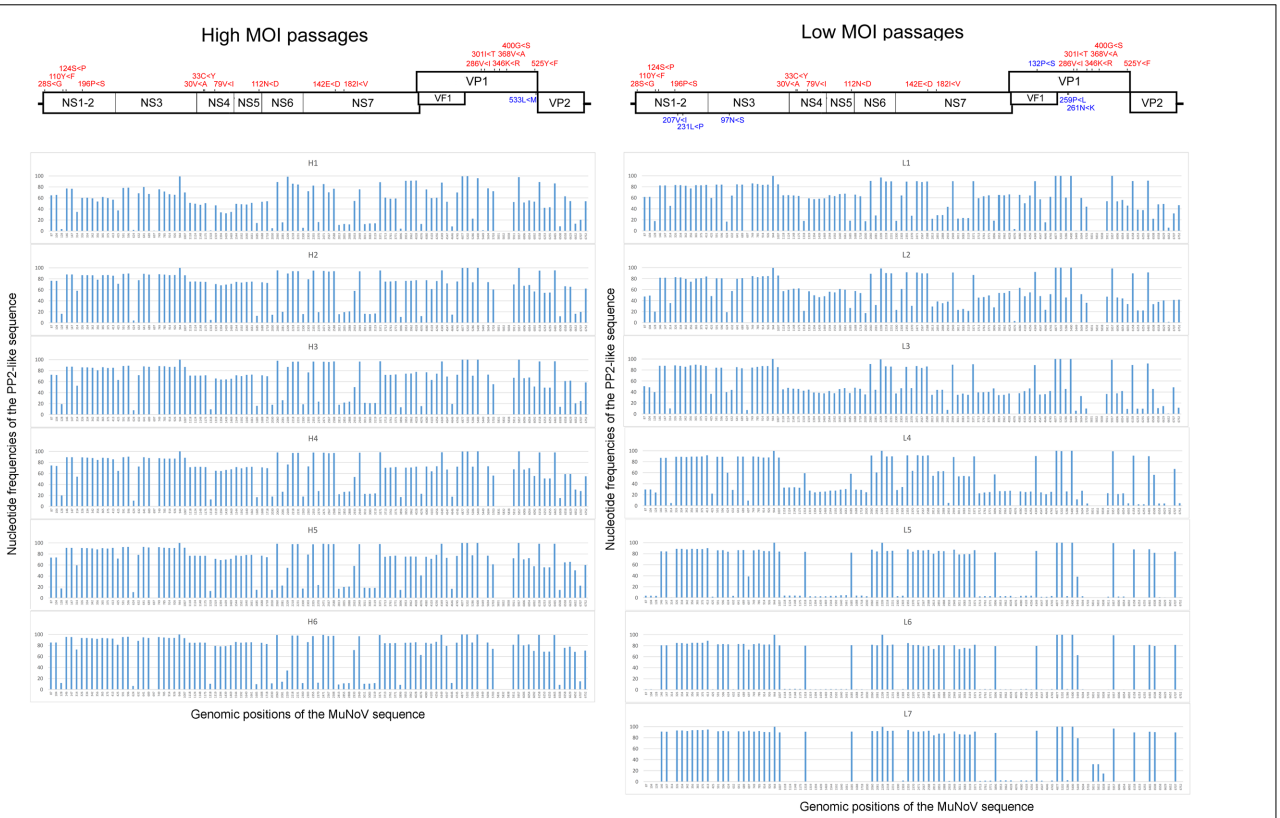


FIGURE 2 | Frequencies of nucleotide differences between the PP2 and PP3 sequences during MuNoV cell passage. Changes in PP2-sequence frequencies during high- and low-MOI passages (**Left** and **Right**, respectively). The horizontal axis indicates the genomic position of the sequences, and the vertical axis indicates the nucleotide frequencies of the PP2-like sequence. Upper panels depict schematics of the MuNoV genome, with indications of non-synonymous nucleotide changes translated to amino acids. Numbers indicate amino acid position affected by the nucleotide variations, and capital letters flanking "<" represent amino acids encoded by PP3 or PP2 sequences, respectively. Letters in red are common changes observed between the high- and low-MOI passages, and those in blue represent changes unique between passages.

0.899 and 0.334 for L7 using the conservative and default settings, respectively (**Table 1**). For the ORF3 region, haplotype reconstruction using the conservative settings generated single major haplotypes for the final passages at both MOIs and major haplotype frequencies of 0.688 and 0.924 for the high- and low-MOIs using the default setting (**Table 1**).

The major haplotypes identified in the full genome sequences from the final high- and low-MOI passages were distinct—with 81 and 52 nucleotide differences from the original PP3 sequence, respectively, and 35 of which were shared (**Figure 3**). Those changes consisted of 16 non-synonymous and 13 missense mutations, including 8 shared (**Figure 3**) and did not seem localized to a particular region in the viral genome. Notably, the major haplotype sequences from the final passages generated from the full genome and individual ORFs were identical. Moreover, the major haplotypes for ORF2 and ORF3 in the final passages generated by default and conservative settings were identical.

Close examination of major ORF2 haplotypes in low MOI passage with the default settings revealed dynamic changes in number and frequency. As such, we examined the frequencies of 15 haplotypes that accounted for >3% of those observed in any single passage and found temporal shifts for several haplotypes that were distinct from the majority haplotypes in the final passage (**Figure 4A**). Similar changes in ORF3 haplotype frequencies were also noted (**Figure 4B**).

Genetic variants present in the majority of the viral population in the final passages represented those better adapted for propagation in RAW264.7 cells. Conversely, variants present in earlier passages that were not observed in later passages were likely unfit for growth in RAW264.7 cells. Therefore, viruses harboring PP2-like sequences may represent viral populations not adapted to *in vitro* propagation. We cloned the PP2-like sequences from viral RNA extracted from the PP3 virus and generated recombinant viruses; however, the cloned sequences only covered a minor fraction of the PP2 nucleotide variations, and the recombinant viruses showed no growth defects in RAW264.7 cells (Supplemental Presentation 2).

To determine the cumulative effect of all PP2 nucleotide variations on viral propagation in RAW264.7 cells, we synthesized a PP2 genome that incorporated all nucleotide variations that differed from PP3 detected by deep sequencing (Supplemental Table 4). The PP2com genome was inserted into a PP3 construct previously shown to produce infectious particles

TABLE 1 | Haplotype frequencies generated by QuasiRecomb.

(A)

	High-MOI passages						Low-MOI passages						
	H1	**H2**	**H3**	**H4**	**H5**	**H6**	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**
Full genome													
Average coverage	847	5975	2380	2058	3792	2883	3056	223	227	2295	2506	4032	3148
No. of haplotypes	9966	16	2903	4176	12	29	34	9740	8575	87	1	8	2
Freq. >3%	0	6	0	0	4	5	7	0	0	12	1	2	2
Freq. (1st)	0.0003	0.317	0.009	0.008	0.313	0.54	0.464	0.004	0.001	0.258	1	0.923	0.908
ORF1													
Average coverage	871	6018	2464	2098	3829	2950	3266	234	240	2382	2636	4233	3153
No. of haplotypes	8561	8	170	611	10	2	20	5973	5026	54	1	2	1
Freq. >3%	0	7	5	6	6	2	7	0	0	2	1	2	1
Freq. (1st)	0.001	0.438	0.424	0.527	0.446	0.787	0.372	0.002	0.003	0.681	1	0.715	1
ORF2													
Average coverage	870	6407	2450	2183	4073	3093	2772	201	201	2118	2303	3728	3217
No. of haplotypes	64	4	215	222	4	15	501	124	8	1	1	1	2
Freq. >3%	6	2	11	7	2	5	0	10	7	1	1	1	2
Freq. (1st)	0.044	0.65	0.06	0.051	0.658	0.691	0.023	0.067	0.379	1	1	1	0.899
ORF3													
Average coverage	716	5189	1780	1643	3194	2131	2434	218	215	2284	2269	3662	3243
No. of haplotypes	1	1	1	1	1	1	1	1	2	1	1	1	1
Freq. >3%	1	1	1	1	1	1	1	1	2	1	1	1	1
Freq. (1st)	1	1	1	1	1	1	1	1	0.556	1	1	1	1

(B)

	High-MOI passages						Low-MOI passages						
	H1	**H2**	**H3**	**H4**	**H5**	**H6**	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**
ORF2													
Average coverage	870	6407	2450	2183	4073	3093	2,772	201	201	2,118	2,303	3,728	3,217
No. of haplotypes	3569	3382	4118	4455	3978	1623	5343	3976	1876	972	131	88	117
Freq. >3%	0	0	0	0	0	1	0	0	0	4	6	7	6
Freq. (1st)	0.004	0.008	0.004	0.004	0.005	0.059	0.002	0.003	0.023	0.101	0.377	0.364	0.334
ORF3													
Average coverage	716	5189	1780	1643	3194	2131	2434	218	215	2284	2269	3662	3243
No. of haplotypes	4	4	4	4	4	4	8	4	4	4	4	4	2
Freq. >3%	4	4	4	4	4	3	4	4	2	3	3	3	2
Freq. (1st)	0.486	0.555	0.49	0.421	0.509	0.688	0.369	0.369	0.5	0.67	0.82	0.738	0.924

*QuasiRecomb was used to generate haplotypes of entire genome (Full genome) and individual ORFs, ORF1, ORF2, and ORF3, under conservative setting in **(A)**, and of ORF2 and ORF3 under default setting in **(B)**. Number of average reads of the sequencing data (Average coverage), number of haplotypes (No. of haplotypes), number of haplotypes accounted higher than 3% (Freq. >3%), and frequency of the major haplotype (Freq. 1st) from each passage are shown.*

upon transfection into 293T cells. Notably, the PP2com viral genome failed to produce infectious particles (**Figure 5B**,**b**), whereas the PP3 sequence generated high viral titers infectious to RAW264.7 cells (**Figure 5B**,**a**). Further, conversion of the PP2com ORF1 or ORF23 region to the PP3 sequence was unable to restore particle production (**Figures 5B**,**c**,**d**). Subsequent analysis of ORF1 protein expression revealed that NS1-2 and NS7 were detectable in cells transfected with the PP2com sequence (**Figure 5C**); however, VP1 expression was undetectable, given that the extent of sub-genomic replication is low when using the plasmid-based reverse-genetics system (Katayama et al., 2014). Examination of interactions between the PP2com and PP3 sequences showed that the PP2com sequence interfered with the production of infectious particles from the PP3 sequence, as co-transfection of the PP3 and PP2com genomes yielded no infection (**Figure 5B**,**e**). However, presence of the PP2com ORF1 region alone was insufficient to block the PP3-mediated virus production in one experiment (**Figure 5B**,**f**), but in two other experiments infectious virus was not produced from PP3 genome with the presence of ORF1 PP2 genome. In contrast, co-transfection of the PP3 genome and ORF23 PP2 generated infectious particles (**Figure 5B**,**g**), as did that of

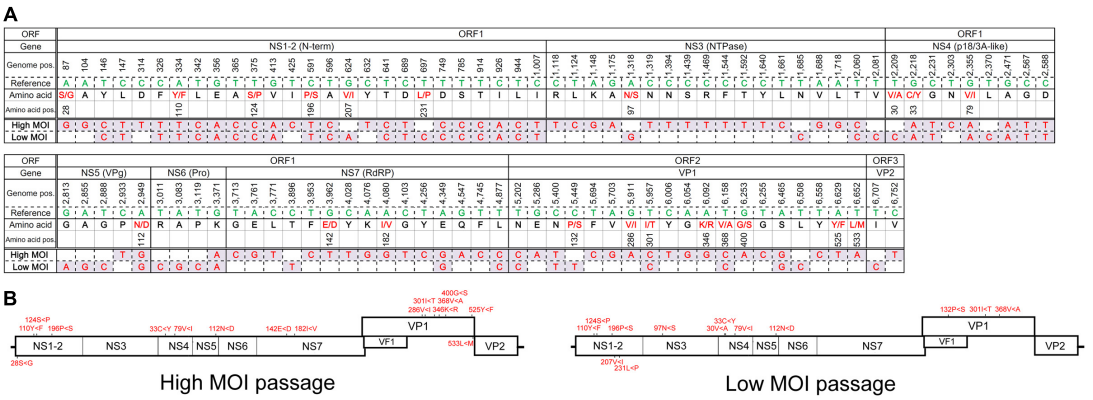


FIGURE 3 | Major haplotype sequences found following the final high- or low-MOI passage. **(A)** Sequence of the major haplotype following the final high- and low-MOI passages. Nucleotides that differed from those found in the PP3 (Reference) sequence are shown along with their genome position (Genome pos.) and the encoded amino acid residue (Amino acid) and their position (Amino acid pos.) in the PP3- or PP2-encoded protein, respectively. Non-synonymous changes are depicted in red. **(B)** Non-synonymous changes of each major haplotype are depicted in the MuNoV genome. Amino acid changes are described similar to those in **Figure 2**.

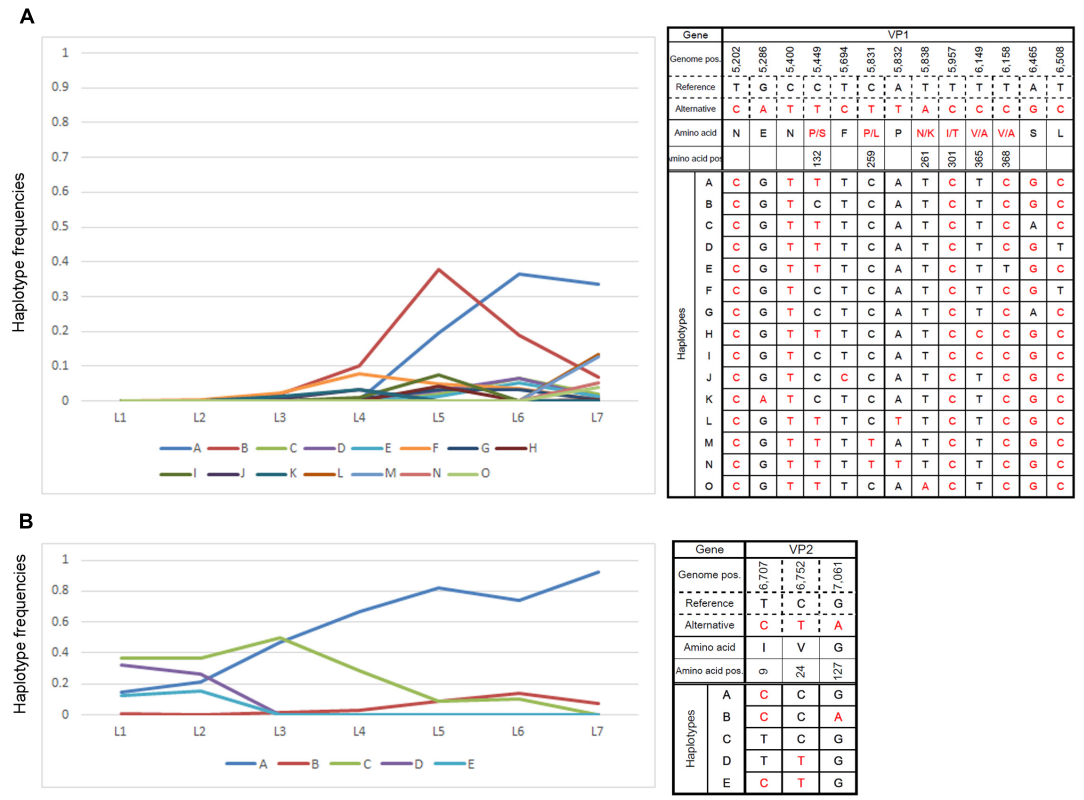


FIGURE 4 | Changes in major haplotype frequencies during each passage. Changes in haplotype frequencies in **(A)** VP1- and **(B)** VP2-encoding regions during low-MOI passages were examined by QuasiRecomb using the default setting. Fifteen haplotypes accounted for >3% of the changes observed in any of the passages were selected and monitored for the frequency of nucleotide changes observed during each passage. The right panels indicate the haplotype sequence along with genome position and possible amino acid change. The left panels show changes in nucleotide frequencies of the selected haplotypes as a line graph. The vertical and horizontal axes indicate the haplotype frequencies and samples for the respective cell passage, respectively.

ORF1 PP3 and ORF23 PP2 (**Figure 5B,h**). The quantification of viral particles in the supernatant of cells transfected with PP3, PP3 + ORF23 PP2, and ORF1 PP2 + ORF23 PP2 contained approximately 2.45 × 10^3/mL, 2.67 × 10^4/mL, and 8.09 × 10^3/mL CCID50, respectively, whereas no particles were found in supernatants from cells transfected with PP3 and ORF1

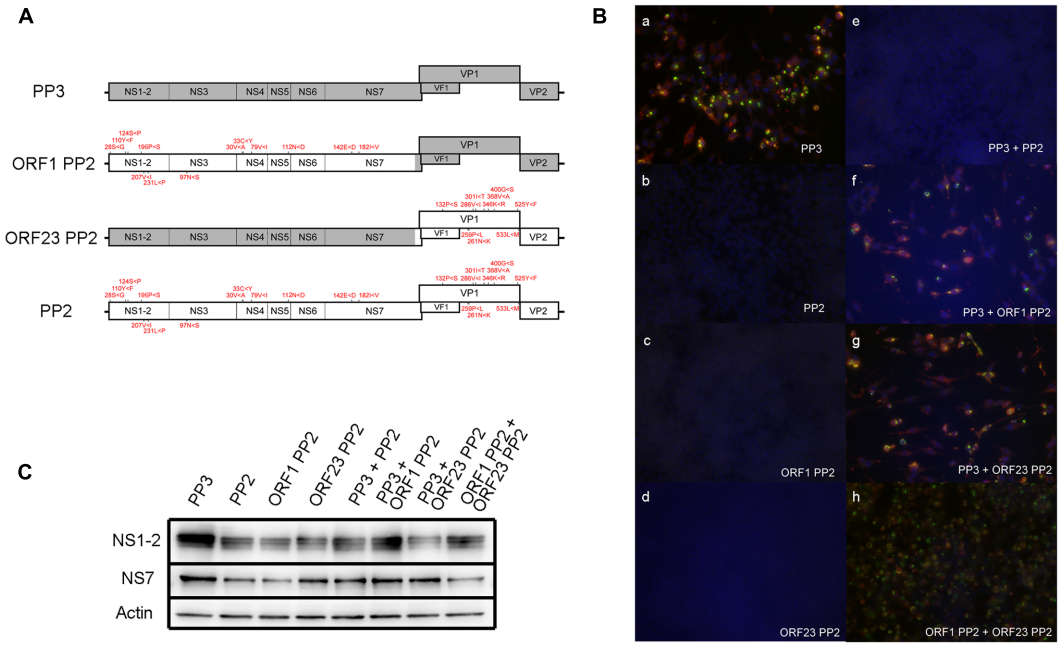


FIGURE 5 | Examination of the ability of the PP2com genome to generate infectious particles. **(A)** Schematic of the recombinant MuNoV genome depicting regions containing PP3 and PP2com sequences. The amino acid residues encoded by the PP2com genome that differed from those encoded by the PP3 genome are indicated in red. **(B)** Supernatant of transfected cells in **(C)** was used to infect RAW264.7 cells. At 48 hpi, cells were fixed and detected for VPg (Green) and VP1 (red) by immunofluorescent detection with staining for DNA by DAPI (Blue) (see "Immunofluorescent Staining" in Materials and Methods). **(C)** Detection of the ORF1 products NS1-2, NS7, and actin by Western blot from the lysates of 293T cells transfected with pMNVs expressing the respective genome. The first four sets of samples were prepared by cotransfecting 2 μg of pT7 MNV S7 and 2 μg each of plasmid expressing the designated recombinant MuNoV genome. The last four sets of the samples were prepared by transfecting the plasmid in combination with those expressing the respective genome.

PP2 cDNA. Collectively, these results indicated that the presence of PP2-like nucleotide variations attenuated viral growth in RAW264.7 cells and interfered with the production of infectious particles from fully adapted sequences. Moreover, variants in the ORF1 region exhibited a *trans*-dominant inhibitory effect; however, co-expression of ORF1 PP2 and ORF23 PP2 to 293T cells generated particles infectious to RAW264.7 cells, indicating that interaction between the PP2com and PP3 sequence was complex—showing both complementation and interference—and dependent upon the context of the sequences.

DISCUSSION

In this study, we examined details associated with changes in nucleotide frequencies in the MuNoV genome during viral passage following infection with two different MOIs. Interestingly, the number of variants decreased over time depending on the culture conditions, with greater variations generally observed with high-MOI infection. Haplotype analysis revealed that the major haplotype sequences in the final passages differed between high- and low-MOI cultures, suggesting that the initial viral populations contained multiple cell-adaptable sequences. Moreover, the genetic variations lost during in-cell propagation were assembled into a single genome and used to transfect cells, but was unable to produce infectious particles itself, as well as those from cell-adapted PP3. Although the key nucleotide changes dictating cell adaptation in the MuNoV S7 strain have yet to be elucidated and will be examined in future work, our results revealed the elaborate interplay among viral variants to select the sequences better-adapted to propagation in cell culture. To elaborate, our experimental procedures are depicted in **Figure 1**.

Murine norovirus sequences among the known strains are relatively less divergent with mostly fixed genome lengths (Thackray et al., 2007); however, variations in the MuNoV S7 strain exhibit large diversity, indicating the coexistence of multiple haplotypes in the viral isolate. When compared with known nucleotide variations in the MNV-1 (Wobus et al., 2004) and MNV NIH2409 strains (Barron et al., 2011) during cell passage, only two synonymous variations at +1409 and +2081 in MNV-1 were found in the S7 strain (Wobus et al., 2004; Bailey et al., 2008), although no changes associated with enhanced viral growth were detected in this study. Instead, we discovered nearly 100 nucleotide variations—including ∼20 non-synonymous changes, the frequency of which dynamically altered during the viral passages. The nucleotide changes seemed evenly distributed throughout the genome with slight emphasis at the anterior coding region of ORF1 and very few in ORF3 and 3′UTR regions, which contrasted to the results by others reporting high degree of variation at the ORF3 region (Mauroy et al., 2016).

Given the limited viability of the PP2com genome (**Figure 5**), these findings indicated that the processes associated with cell adaptation involved multiple rounds of nucleotide alteration or selection.

In this study, we selected viral passages from cells infected at two different MOIs. High-MOI passage might allow for the maintenance of sequence variations and avoid sequence convergence or opportunistic genetic drift that result from genetic bottlenecks because of small subsets in the viral population (Andino and Domingo, 2015). Even under such conditions, haplotype analysis indicated that a single haplotype could flourish to predominate the viral population (**Figure 3**), suggesting the presence of selective pressure for variants better suited for propagation in culture. Not surprisingly, low-MOI passage assumed conditions that resembled a genetic bottleneck wherein the majority of cells do not assume infection with multiple viral clones. Although the secondary infection over the following 48 h period did not warrant conditions limiting entry of infectious clones, sequences exhibited convergence during early phase passages (**Figure 3**) and showed changes in haplotype frequency indicative of genetic drift (**Figure 4**). Genetic drift depends on the diversity in the founder viral population and the number of host cells. Since approximately 1×10^6 cells were used in the present study, the viral populations selected during passage would be no less than 1×10^6. Thus, our observation of viral sequence convergence to a single haplotype likely resulted from selective pressure rather than genetic drift, where sequence selection occurs by chance.

The two different passages generated distinct haplotype sequences that predominated the final viral populations. It is possible that the major haplotypes observed in the low-MOI passage were the product of genetic drift and could differ with prolonged passage. Considering that two distinct haplotypes were observed during the two different passages, and since serial passage of the P3 virus generated a major haplotype identical to PP3, our results suggested the presence of multiple viral haplotypes in the original viral isolate or founder population.

Assuming that PP3 was among the most cell-adapted sequences, we constructed the PP2com genome using sequence variations lost during cell passage. As expected, the PP2com genome was unable to produce infectious particles upon expression using the reverse-genetic system (**Figure 5**). This was not attributed to defects in viral gene expression, as the protein products encoded by ORF1 were expressed at levels comparable with those of PP3. Moreover, the co-transfection of constructs expressing ORF1 PP2 and ORF23 PP2 generated viral particles infectious to RAW264.7 cells (**Figures 5B,C**). Interestingly, co-expression of the PP2com and PP3 genomes blocked the production of infectious particles from PP3, indicating the presence of sequence elements in PP2com that interfered with viral growth *in trans*. Further, presence of the PP2com ORF1 region blocked viral particle production from PP3 genome, but the PP2com ORF23 region only showed a marginal blockade (**Figure 5B**,**g**). Such effect could be an accumulative since ORF1 and ORF23 from PP2com harbored 81 and 24 nucleotide differences from PP3 sequence, respectively, thus the number of nucleotide differences could be positively correlated with the strength of interference. Such genetic interference could also depend upon the sequence context as co-expression of the ORF1 and ORF23 PP2 genomes generated infectious particles despite that the PP2com-derived ORF1 sequence was present (**Figure 5B**,**h**). In addition, the coexistence of multiple viral haplotypes has been shown to generate complex interplay, including interference irrespective of coding or non-coding sequences (Ojosnegros et al., 2010; Acevedo et al., 2014). Thus, presence of PP2com-like sequences in the founder viral population would likely impart negative effects on viral propagation *in vitro*.

Here, we used deep sequencing to reveal that the frequencies of MuNoV S7 variants dynamically changed to converge to a cell-adaptable viral sequence during multiple viral passages under cell-culture conditions. The selective process associated with passages of MuNoV S7 in the cultured cells was more complex than previously reported in MNV-1, where only two nucleotide changes were sufficient (Wobus et al., 2004; Bailey et al., 2008). Our findings that the artificial PP2com genome exhibited *trans*-dominant negative effects on viral growth against a cell-adapted genome suggested elaborate interplay among viral variants during the course of the cell-adaptation process in MuNoV. These findings enhanced the understanding of processes related to laboratory adaptation of non-cultivable viruses, especially to those without relevant cell culture systems.

AUTHOR CONTRIBUTIONS

KaK and AN designed the study, performed the experiments, and wrote the manuscript. HT, AK, KuK, and RT-T performed the experiments. TK and KY performed bioinformatics analyses. All authors critically revised the manuscript and approved the final version.

FUNDING

This study was supported in part by funding from a commissioned project for the Research on Emerging and Re-emerging Infectious Diseases from the Japanese Ministry of Health, Labor, and Welfare, the Research Program on Emerging and Re-emerging Infectious Diseases from the Japan Agency for Medical Research and Development program, and the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research, KAKENHI.

ACKNOWLEDGMENT

We thank Dr. Yukinobu Tohya (Nihon University) for providing the MuNoV S7 strain.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.01091/full#supplementary-material

REFERENCES

Acevedo, A., Brodsky, L., and Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505, 686–690. doi: 10.1038/nature12861

Andino, R., and Domingo, E. (2015). Viral quasispecies. *Virology* 47, 46–51. doi: 10.1016/j.virol.2015.03.022

Arias, A., Urena, L., Thorne, L., Yunus, M. A., and Goodfellow, I. (2012). Reverse genetics mediated recovery of infectious murine norovirus. *J. Vis. Exp.* 64:e4145. doi: 10.3791/4145

Baert, L., Wobus, C. E., Van Coillie, E., Thackray, L. B., Debevere, J., and Uyttendaele, M. (2008). Detection of murine norovirus 1 by using plaque assay, transfection assay, and real-time reverse transcription-PCR before and after heat exposure. *Appl. Environ. Microbiol.* 74, 543–546. doi: 10.1128/AEM.01039-07

Bailey, D., Thackray, L. B., and Goodfellow, I. G. (2008). A single amino acid substitution in the murine norovirus capsid protein is sufficient for attenuation in vivo. *J. Virol.* 82, 7725–7728. doi: 10.1128/JVI.00237-08

Barron, E. L., Sosnovtsev, S. V., Bok, K., Prikhodko, V., Sandoval-Jaime, C., Rhodes, C. R., et al. (2011). Diversity of murine norovirus strains isolated from asymptomatic mice of different genetic backgrounds within a single U.S. research institute. *PLoS ONE* 6:e21435. doi: 10.1371/journal.pone.0021435

Duizer, E., Schwab, K. J., Neill, F. H., Atmar, R. L., Koopmans, M. P., and Estes, M. K. (2004). Laboratory efforts to cultivate noroviruses. *J. Gen. Virol.* 85(Pt 1), 79–87. doi: 10.1099/vir.0.19478-0

Ettayebi, K., Crawford, S. E., Murakami, K., Broughman, J. R., Karandikar, U., Tenge, V. R., et al. (2016). Replication of human noroviruses in stem cell-derived human enteroids. *Science* 353, 1387–1393. doi: 10.1126/science.aaf5211

Fischer, A. E., Wu, S. K., Proescher, J. B., Rotem, A., Chang, C. B., Zhang, H., et al. (2015). A high-throughput drop microfluidic system for virus culture and analysis. *J. Virol. Methods* 213, 111–117. doi: 10.1016/j.jviromet.2014.12.003

Haga, K., Fujimoto, A., Takai-Todaka, R., Miki, M., Doan, Y. H., Murakami, K., et al. (2016). Functional receptor molecules CD300lf and CD300ld within the CD300 family enable murine noroviruses to infect cells. *Proc. Natl. Acad. Sci. U.S.A.* 113, E6248–E6255. doi: 10.1073/pnas.1605575113

Jones, M. K., Watanabe, M., Zhu, S., Graves, C. L., Keyes, L. R., Grau, K. R., et al. (2014). Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* 346, 755–759. doi: 10.1126/science.1257147

Karst, S. M., Wobus, C. E., Lay, M., Davidson, J., and Virgin, H. W. T. (2003). STAT1-dependent innate immunity to a Norwalk-like virus. *Science* 299, 1575–1578. doi: 10.1126/science.1077905

Katayama, K., Murakami, K., Sharp, T. M., Guix, S., Oka, T., Takai-Todaka, R., et al. (2014). Plasmid-based human norovirus reverse genetics system produces reporter-tagged progeny virus containing infectious genomic RNA. *Proc. Natl. Acad. Sci. U.S.A.* 111, E4043–E4052. doi: 10.1073/pnas.1415096111

Kitajima, M., Oka, T., Tohya, Y., Katayama, H., Takeda, N., and Katayama, K. (2009). Development of a broadly reactive nested reverse transcription-PCR assay to detect murine noroviruses, and investigation of the prevalence of murine noroviruses in laboratory mice in Japan. *Microbiol. Immunol.* 53, 531–534. doi: 10.1111/j.1348-0421.2009.00152.x

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Mauroy, A., Taminiau, B., Nezer, C., Ghurburrun, E., Baurain, D., Daube, G., et al. (2016). High-throughput sequencing analysis reveals the genetic diversity of different regions of the murine norovirus genome during *in vitro* replication. *Arch. Virol.* 162, 1019–1023. doi: 10.1007/s00705-016-3179-y

McFadden, N., Bailey, D., Carrara, G., Benson, A., Chaudhry, Y., Shortland, A., et al. (2011). Norovirus regulation of the innate immune response and apoptosis occurs via the product of the alternative open reading frame 4. *PLoS Pathog.* 7:e1002413. doi: 10.1371/journal.ppat.1002413

Minor, P. D. (2015). Live attenuated vaccines: historical successes and current challenges. *Virology* 47, 379–392. doi: 10.1016/j.virol.2015.03.032

Ojosnegros, S., Beerenwinkel, N., Antal, T., Nowak, M. A., Escarmis, C., and Domingo, E. (2010). Competition-colonization dynamics in an RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2108–2112. doi: 10.1073/pnas.0909787107

Taube, S., Kolawole, A. O., Hohne, M., Wilkinson, J. E., Handley, S. A., Perry, J. W., et al. (2013). A mouse model for human norovirus. *mBio* 4:e450-13. doi: 10.1128/mBio.00450-13

Thackray, L. B., Wobus, C. E., Chachu, K. A., Liu, B., Alegre, E. R., Henderson, K. S., et al. (2007). Murine noroviruses comprising a single genogroup exhibit biological diversity despite limited sequence divergence. *J. Virol.* 81, 10460–10473. doi: 10.1128/JVI.00783-07

Topfer, A., Zagordi, O., Prabhakaran, S., Roth, V., Halperin, E., and Beerenwinkel, N. (2013). Probabilistic inference of viral quasispecies subject to recombination. *J. Comput. Biol.* 20, 113–123. doi: 10.1089/cmb.2012.0232

Wobus, C. E., Karst, S. M., Thackray, L. B., Chang, K. O., Sosnovtsev, S. V., Belliot, G., et al. (2004). Replication of norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biol.* 2:e432. doi: 10.1371/journal.pbio.0020432

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kitamoto, Takai-Todaka, Kato, Kanamori, Takagi, Yoshida, Katayama and Nakanishi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 March 2017
doi: 10.3389/fmicb.2017.00410

Evolutionary Constraints on the Norovirus Pandemic Variant GII.4_2006b over the Five-Year Persistence in Japan

Hironori Sato[1], Masaru Yokoyama[1], Hiromi Nakamura[1], Tomoichiro Oka[2], Kazuhiko Katayama[2,3], Naokazu Takeda[4,5], Mamoru Noda[6], Tomoyuki Tanaka[7] and Kazushi Motomura[1,4,5]*

[1] Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan, [2] Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan, [3] Graduate School of Infection Control Sciences, Kitasato University, Tokyo, Japan, [4] Research Institute for Microbial Diseases, Osaka University, Osaka, Japan, [5] Thailand-Japan Research Collaboration Center on Emerging and Re-emerging Infections, Nonthaburi, Thailand, [6] National Institute of Health Sciences, Tokyo, Japan, [7] Sakai City Institute of Public Health, Osaka, Japan

Norovirus GII.4 is a major cause of global outbreaks of viral gastroenteritis in humans, and has evolved by antigenic changes under the constantly changing human herd immunity. Major shift in the pandemic GII.4 strain periodically occurs concomitant with changes in the antigenic capsid protein VP1. However, how the newly emerged strain evolves after the onset of pandemic remains unclear. To address this issue, we examined molecular evolution of a pandemic lineage, termed the GII.4_2006b, by using the full-length viral genome and VP1 sequences ($n = 317$) from stools collected at 20 sites in Japan between 2006 and 2011. Phylogenetic tree showed a radial diversification of the genome sequences of GII.4_2006b, suggesting a rapid genetic diversification of the GII.4_2006b population from a few ancestral variants. Impressively, amino acid sequences of the variable VP1 in given seasons remained as homogeneous as those of viral enzymes under annual increase in the nucleotide diversity in the VP1 coding region. The Hamming distances between the earliest and subsequent variants indicate strong constraints on amino acid changes even for the highly variable P2 subdomain. These results show the presence of evolutionary constraints on the VP1 protein and viral enzymes, and suggest that these proteins gain near maximal levels of fitness benefits in humans around the onset of the outbreaks. These findings have implications for our understanding of molecular evolution, mechanisms of the periodic shifts in the pandemic NoV GII.4 strains, and control of the NoV GII.4 pandemic strain.

Keywords: norovirus, pandemic strain, GII.4, capsid, molecular evolution, evolutionary constraints

OPEN ACCESS

Edited by:
Stefan Taube,
University of Lübeck, Germany

Reviewed by:
Hirotaka Ode,
National Hospital Organization Nagoya Medical Center, Japan
Janet Mans,
University of Pretoria, South Africa

****Correspondence:***
Hironori Sato
hirosato@nih.go.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *06 January 2017*
Accepted: *27 February 2017*
Published: *13 March 2017*

Citation:
Sato H, Yokoyama M, Nakamura H, Oka T, Katayama K, Takeda N, Noda M, Tanaka T and Motomura K (2017) Evolutionary Constraints on the Norovirus Pandemic Variant GII.4_2006b over the Five-Year Persistence in Japan. Front. Microbiol. 8:410. doi: 10.3389/fmicb.2017.00410

INTRODUCTION

Norovirus (NoV) is a non-enveloped RNA virus that belongs to the family *Caliciviridae*. NoV is divided into multiple genogroups, which are further subgrouped into more than 40 genotypes (Zheng et al., 2006; Donaldson et al., 2010; Robilotti et al., 2015). Among them, genogroup II genotype 4 (GII.4) is especially important in public health, because it is the leading cause of NoV-associated acute gastroenteritis in humans (Noel et al., 1999; Siebenga et al., 2007;

Zheng et al., 2010). NoV is highly variable, and emergence of a novel GII.4 variant is in some cases associated with global outbreaks of gastroenteritis (Lopman et al., 2004; Vipond et al., 2004; Reuter et al., 2005; Bull et al., 2006; Gallimore et al., 2007; Motomura et al., 2008). A major shift in the pandemic strain occurs with antigenic mutations (Lindesmith et al., 2008, 2011, 2012a, 2013) and viral genome recombination at the ORF1/2 boundary region (Motomura et al., 2010; Eden et al., 2013).

The VP1 protein is the major structural protein of the mature virion, which protrudes from the virion surface and plays pivotal roles in the viral interactions with hosts. The VP1 protein is composed of two domains, protruding (P) and shell (S) (Prasad et al., 1999). The P domain is further divided into two subdomains, P1 and P2 (Prasad et al., 1999). The P2 subdomain is placed on the tip of the VP1 protein and constitutes the major antigenic site around the binding site to the putative receptor(s) for infection (Donaldson et al., 2010). This structural feature causes sequence variation (Lindesmith et al., 2008, 2011, 2012a, 2013; Bok et al., 2009; Debbink et al., 2012) and structural diversity (Chen et al., 2004, 2006; Donaldson et al., 2010), particularly in the P2 subdomain. Meanwhile, the functional importance of the P2 subdomain can cause suppression of deleterious changes and/or changes that reduce viral replication fitness. However, very little is known about evolution of the VP1 protein during viral maintenance in human populations.

To address this issue, we examined here molecular evolution of the VP1 protein of a pandemic lineage, termed GII.4_2006b, which is also known as the GII.4 Den Haag 2006b. In the autumn/winter of 2006, the national epidemiological surveillance of infectious diseases in Japan reported an unusual increase in the number of outbreaks of NoV infections (Infectious Disease Surveillance Center[1]). This augmentation was associated with the nationwide spread of a newly emerging GII.4 variant (Motomura et al., 2008), termed GII.4_2006b. The GII.4_2006b initially coexisted as a minority strain among various other NoV lineages in Japan, but starting in October of 2006 it spread extremely rapidly and remained as the major epidemic variant across Japan between 2006 and 2009 (Motomura et al., 2008, 2010). In this study, we characterized nucleotide and amino acid diversities of the VP1 proteins, using serially collected full-length 317 NoV genome and VP1 sequences from infections in Japan between 2006 and 2011. The obtained results show long-term persistent of GII.4_2006b in human populations in Japan as a dominant GII.4 subpopulation. Interestingly, both the VP1 protein and viral enzymes had remained as highly homogeneous populations, indicating strong evolutionary constraints on changes in these proteins following the onset of the outbreaks.

MATERIALS AND METHODS

NoV Genome Sequencing

Stool specimens were collected from individuals with acute gastroenteritis at 20 regional public health institutes in Japan between May 2006 and March 2011 in compliance with the Food Sanitation Law of Japan, according to the methods for the protection of personal information (including methods for anonymization in an unlinkable fashion). The research was approved by research and ethical committee in National Institute of Infectious Diseases. Three to five stool specimens were collected at each site in each year. NoV genome sequences were obtained from the stool specimens as described previously (Motomura et al., 2008, 2010).

Genotype Determination

Norovirus genotype was determined by construction of phylogenetic trees of viral genome sequences. Multiple sequence alignments were done as described previously (Motomura et al., 2008, 2010) using the MAFFT (Katoh et al., 2009) and alignment tools implemented in the MEGA software suite (Tamura et al., 2011). Phylogenetic trees were constructed as described previously (Motomura et al., 2008, 2010) using MEGA software (Tamura et al., 2011). The reliability of interior branches in the tree was assessed by the bootstrap method with 1,000 resamplings.

Analysis of Diversity of Sequence Population

Mean diversity in the entire sequence population was computed with the "*Sequence Diversity*" menu in MEGA software suite (Tamura et al., 2011). The overall pairwise mean distance between the sequences was computed with the "*Distances*" menu in MEGA. As substitution models, a maximum composite likelihood and a Poisson model were used for nucleotide and amino acid sequences, respectively. Variance was estimated by the bootstrap method with 100 to 500 bootstrap replications.

Analysis of Individual Amino Acid Variation

Amino acid variations at each position of the VP1 (1–530) were calculated as previously described with a multiple sequence alignment as described previously for other viral proteins (Naganawa et al., 2008; Oka et al., 2009; Takahata et al., 2017) on the basis of Shannon's equation (Shannon, 1997):

$$H(i) = -\sum_{x_i} p(x_i) \log_2 p(x_i)$$

$$(x_i = \text{G, A, I, V, ...}), \quad (1)$$

where $H(i)$, $p(x_i)$, and i indicate the amino acid entropy score of a given position, the probability of occurrence of a given amino acid at the position, and the number of positions, respectively. An $H(i)$ score of zero indicates absolute conservation, whereas 4.4 bits for amino acids or 2.0 bits for nucleic acids indicates complete randomness.

Analysis of Amino Acid Substitutions by Hamming Distance

We used the Hamming distance to assess the changeability of the earliest NoV GII.4_2006b variant in Japan. In information

[1]http://idsc.nih.go.jp/iasr/prompt/graph-ke.html

theory, the Hamming distance between two sequences indicates the minimum number of "amino acid substitutions" required to change one sequence into the other. Because the length of amino acid sequences of the VP1 proteins of the NoV GII.4_2006b subpopulations were identical (540 amino acid residues), the Hamming distance measured in this study means the number of different amino acid residues between the earliest and subsequently emerged variants in two aligned sequences. Python was used as the programming language to compute Hamming distances. Hamming distances between the sequence in May 2006 (accession number AB447443; earliest GII.4_2006b sequence in our NoV genome dataset) and the later GII.4_2006b sequences ($n = 249$) were computed by creating a sequence that assigns mismatches and matches at corresponding positions in the two sequences, and then by counting the numbers of the mismatches.

Nucleotide Accession Numbers

The DDBJ database accession numbers of the 317 NoV GII.4 genome sequences used in this study are provided in Supplementary Table S1 ($n = 250$, GII.4_2006b) and Supplementary Table S2 ($n = 67$, GII.4 non-2006b).

RESULTS

Persistence and Diversification of NoV GII.4_2006b Genome in Japan between 2006 and 2011

We obtained 317 genome sequences of NoV GII.4 from the stool specimens collected at 20 sites in Japan between 2006 and 2011 (**Figure 1A**). Eight distinct lineages of NoV GII.4 were

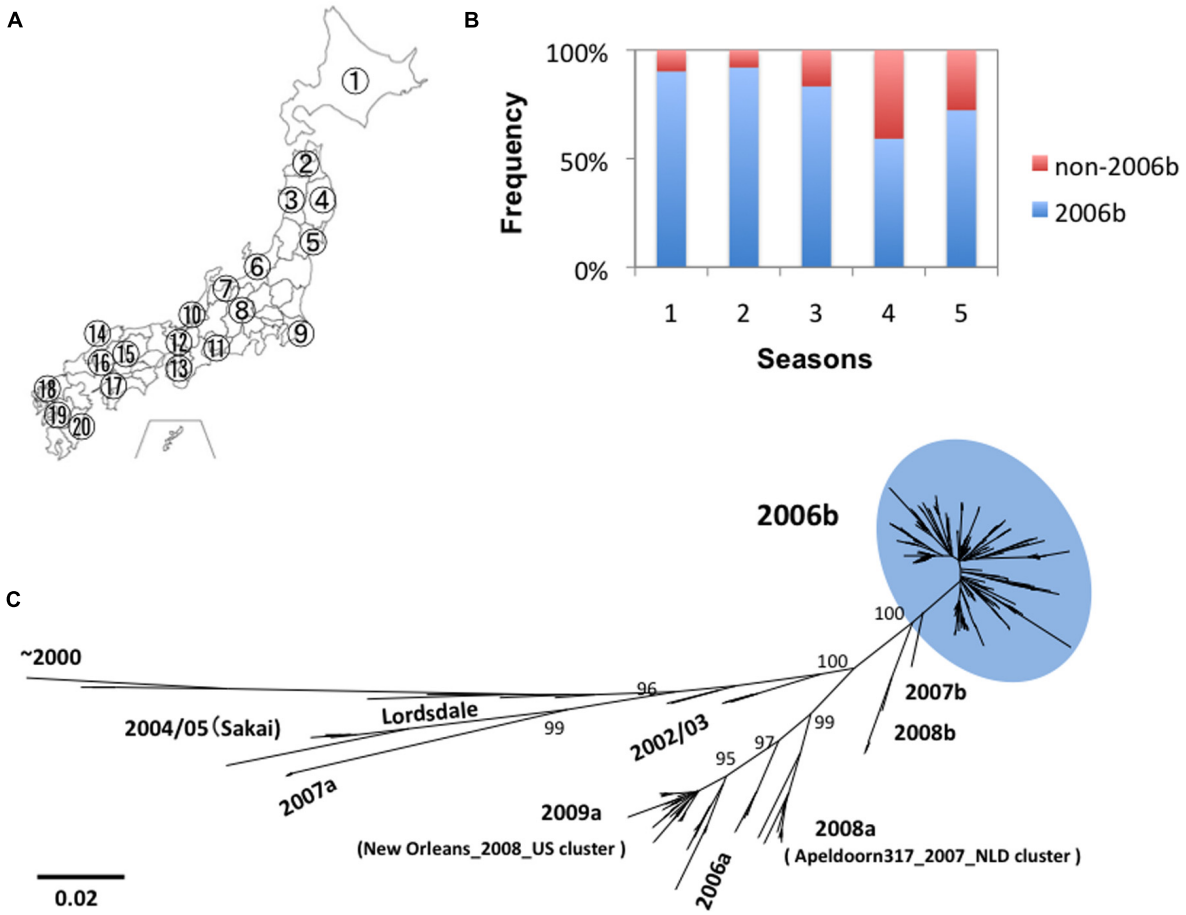


FIGURE 1 | Persistence and diversification of NoV GII.4_2006b genome in Japan between 2006 and 2011. (A) Geographic locations of the 20 sample collection sites in Japan (see Supplementary Tables S1, S2 for the seasons and sites of collection). **(B)** Detection frequency of the GII.4_2006b variants in each period. NoV GII.4 genome sequences ($n = 317$) from stool specimens collected at the 20 sites between May 2006 and March 2011 were divided into five groups on the basis of the collection periods, i.e., 0–11 months (the season 1, $n = 41$), 12–23 months (the season 2, $n = 75$), 24–35 months (the season 3, $n = 78$), 36–47 months (the season 4, $n = 76$), and 48–58 months (the season 5, $n = 47$) after the first detection of GII.4 in May 2006 in Japan in this study. Genotypes of the GII.4 variants were determined with phylogenetic trees of the whole genome sequences as described previously (Motomura et al., 2008, 2010). The detection frequency of the GII.4_2006b genomes among the total GII.4 genomes in each collection period is shown. **(C)** Phylogenetic classification of the NoV GII.4 genome sequences used in this study. The maximum likelihood tree was constructed with the 317 GII.4 genome sequences (about 7.5 kb). The sequence cluster enclosed by a light-blue oval indicates the GII.4_2006b genomes.

identified in this period. These lineages include "2004/05" related to Sakai/04-179/2005/JP cluster, "2006a" related to Yerseke 2006a cluster, "2006b" related to Den Haag 2006b cluster, "2007a," "2007b," "2008a" related to Apeldoorn317/2007/NL cluster, "2008b," and "2009a" related to New Orleans 1805/2009/USA cluster (Motomura et al., 2008, 2010; Supplementary Tables S1, S2). Among the newly emerged eight GII.4 lineages, only the pandemic variant GII.4_2006b had been detected dominantly and continually throughout Japan (**Figure 1B**). The GII.4_2006b represented about 79% ($n = 250$) of the total GII.4 genomes detected during the 5 years. Phylogenetic analysis shows that the GII.4_2006b genome sequences diverged radially from a few roots (**Figure 1C**). These data suggest genetic bottlenecks followed by a rapid genome diversification of GII.4_2006b between 2006 and 2011.

Diversity of NoV GII.4_2006b ORFs

The GII.4_2006b RNA genome encodes three open reading frames, ORF1, ORF2, and ORF3 (**Figure 2A**). ORF1 encodes viral enzymes and non-structural proteins. ORF2 and ORF3 encodes structural proteins, VP1 and VP2, respectively. We first examined whether the sequence diversity is different among the three ORFs by using the 250 GII.4_2006b genome sequences obtained in this study. The phylogenetic tree and mean diversity in the entire sequence population show that the nucleotide diversity was similar among the three ORFs (**Figure 2B**). In contrast, a marked difference was observed in the diversity of amino acid sequences: the ORF1 and ORF2 amino acid sequences remained significantly less diversified than that of ORF3 (**Figure 2C**). The data suggest the presence of constraints on the amino acid changes of proteins encoded by ORF1 and ORF2.

Temporal Changes in the Sequence Diversity of NoV GII.4_2006b

The GII.4_2006b RNA genome encodes eight viral proteins (**Figure 3A**). Shannon entropy of amino acid sequences of the 2006b ORF1 in the present genome dataset indicates that the potential sites for the internal cleavage of the ORF1 precursor protein were perfectly conserved in amino acid levels [$H(i) = 0/0$] for p48/NTPase (Q/G), NTPase/p22 (Q/G), and VPg/Pro (E/A),

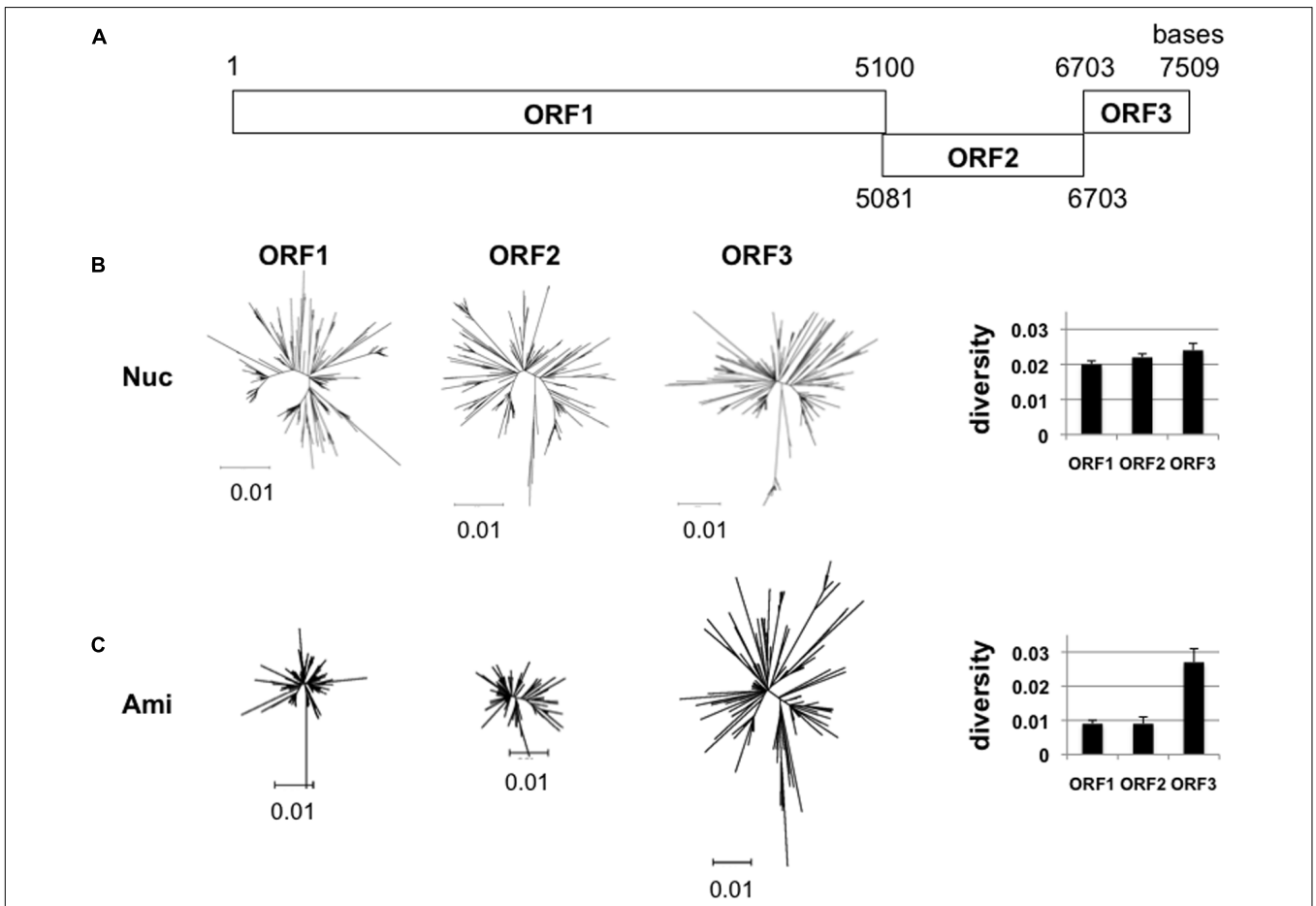


FIGURE 2 | Diversity of NoV GII.4_2006b ORFs. (A) Schematic illustration of the locations of the three ORFs in the NoV GII.4_2006b genome. The nucleotide positions are based on the sequence of the earliest GII.4_2006b in our NoV genome dataset (Supplementary Table S1, Accession No. AB447443). **(B,C)** Neighbor-joining trees (Left) and diversity scores (Right) of the sequences of the ORF1, 2, and 3 regions in the 250 NoV GII.4_2006b genomes identified in **Figure 1C**. The analyses were done with tools included in the MEGA software suite (Tamura et al., 2011). Nucleotide sequences **(B)**. Amino acid sequences **(C)**.

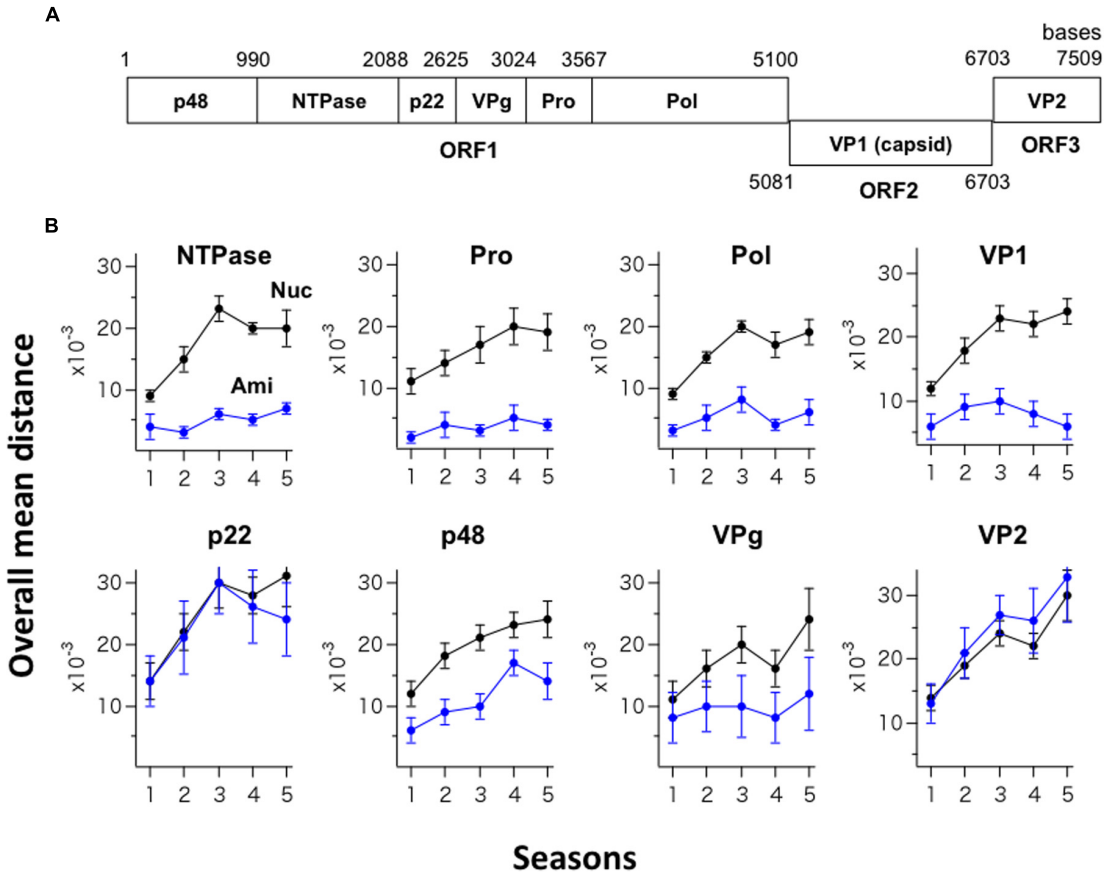


FIGURE 3 | Temporal changes in the sequence diversity of NoV GII.4_2006b. (A) Schematic illustration of the locations of the protein-coding regions in the NoV GII.4_2006b genome. The nucleotide positions are based on the sequence of the earliest GII.4_2006b in our NoV genome dataset (Supplementary Table S1, Accession No. AB447443). **(B)** Temporal changes in dissimilarity of nucleotide and amino acid sequences in protein-coding regions of the GII.4_2006b genome. The 250 sequences of the GII.4_2006b genome were divided into five groups on the basis of the collection periods, i.e., 0–11 months (the season 1, $n = 41$), 12–23 months (the season 2, $n = 75$), 24–35 months (the season 3, $n = 78$), 36–47 months (the season 4, $n = 76$), and 48–58 months (the season 5, $n = 47$) after the first detection of GII.4_2006b in May 2006 in Japan in this study. The overall pairwise mean distance between the sequences was calculated for each season using MEGA software (Tamura et al., 2011). Black and blue circles indicate the mean distances of nucleotide (Nuc) and amino acid (Ami) sequences.

and highly conserved for p22/VPg (E/G) [$H(i) = 0.03/0.08$] and Pro/Pol (E/G) [$H(i) = 0.03/0.03$]. To assess the changeability of individual viral proteins, we examined temporal changes in the sequence diversity of the eight protein-coding regions using the 250 GII.4_2006b genomes. The genome sequences were divided into five groups based on the collected seasons, and the overall mean distance of the sequences in a season was calculated using MEGA. The nucleotide mean distance sequentially increased for the eight protein-coding regions (**Figure 3B**, Nuc), indicating a continuous increase in the dissimilarity of every gene segment in the GII.4_2006b variant population in Japan.

In contrast, the temporal change in amino acid sequence diversity was very different among the eight proteins (**Figure 3B**, Ami). Interestingly, the amino acid mean distance of the generally hypervariable VP1 protein sequences remained comparable to that of three viral enzymes (NTPase, Pro, Pol) for 5 years, with the mean distance remaining at less than 0.01 with small variances (**Figure 3B**, Upper). After the 3rd epidemic season, the VP1 amino acid distance even decreased. Meanwhile, the amino acid mean distances of the p22 and VP2 proteins sharply increased in parallel with an increase in the nucleotide distances (**Figure 3B**, p22 and VP2), suggesting the continuous diversification of these proteins in association with nucleotide diversification. The mean amino acid distance for the p48 protein increased with time yet less extensively than those for the p22 and VP2 proteins (**Figure 3B**, p48). The mean amino acid distance for the VPg protein stayed at relatively low levels with large variances (**Figure 3B**, VPg). In sum, these data suggest the presence of strong constraints on amino acid changes in the capsid protein VP1 and enzymes (NTPase, Pro, Pol) of the GII.4_2006b under the diversification of nucleotide sequences.

Long-term Circulation of the NoV GII.4_2006b Subgroup Carrying the Identical Capsid Protein VP1

We identified a GII.4_2006b subpopulation ($n = 23$) whose nucleotide sequences differed from each other, yet encoded

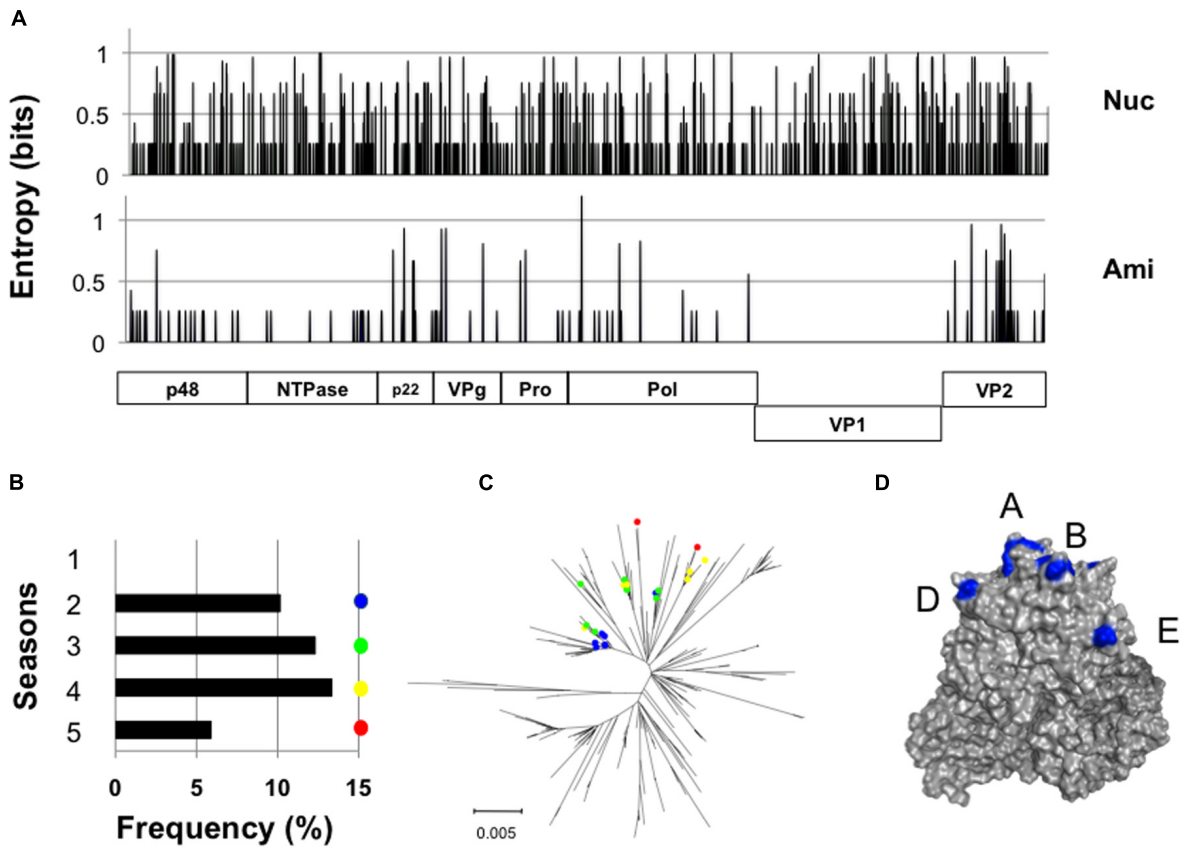


FIGURE 4 | Long-term circulation of the NoV GII.4_2006b subgroup carrying the identical capsid protein VP1. (A) Identification of a GII.4_2006b genome subpopulation group 1 that encodes the identical VP1s in distinct genetic backbones. Nucleotide (Upper) and deduced amino acid (Lower) sequences of the group 1 genomes (n = 23) were aligned using MAFFT software (Katoh et al., 2009), and Shannon entropies at individual positions were calculated as described previously (Naganawa et al., 2008; Oka et al., 2009; Takahata et al., 2017). The distribution of Shannon entropy scores in the GII.4_2006b genome is shown. **(B)** Detection frequency of the group 1 genomes in five seasons between 2006 and 2011 in Japan. **(C)** Neighbor-joining tree of the GII.4_2006b VP1 nucleotide sequences (1620 nucleotides). Colored circles indicate the group 1 sequences. **(D)** P domain dimer model of the GII.4_2006b VP1 protein was constructed as described (Motomura et al., 2008, 2010). Blue residues indicate the GII.4_2006b-specific amino acid substitutions at potential epitopes in the P2 subdomain of the GII.4 VP1 (Lindesmith et al., 2012b).

exactly the same VP1 amino acid sequences (**Figure 4A**). The members of this population, tentatively termed group 1, were detected at distantly located 11 sample collection sites in Japan during the study period (Supplementary Table S1 and Figure S1). They emerged in the second epidemic season in 2007 and continuously circulated without no changes in the VP1 amino acid residues, representing about 6–13% of the GII.4_2006b genomes in each season (**Figure 4B**). The group 1 genomes continuously accumulated nucleotide substitutions in the VP1-coding region (**Figure 4C**), but only synonymous substitutions (**Figure 4A**). The GII.4_2006b variants at the onset of epidemics generally had 10 substitutions at the potential epitopes A, B, D, and E (Lindesmith et al., 2012a) of the VP1 P2 subdomain (P294A, T296S, T298N, A368S, D372E, M333V, R382K, S394T, D407S, and T412N), as compared with the VP1 sequence of the past epidemic variant in 2004/2005 in Japan (Sakai/04-179/2005/JP: accession number BAE98194) (**Figure 4D**). The group 1 VP1 protein had an additional substitution (S393G) on the epitope D.

Temporal Change in Hamming Distance for the NoV GII.4_2006b Capsid Protein VP1

The NoV VP1 protein has an architecture similar to that of the VP1 proteins of other single-stranded RNA viruses (Prasad et al., 1994, 1999; **Figure 5A**). The S domain is highly conserved, whereas the P2 domain is hypervariable among GII.4 variants. To assess the changeability of the P2 domain of the GII.4_2006b, we examined the temporal accumulations of amino acid substitutions in the S, P1, and P2 regions of the GII.4_2006b VP1 using the Hamming distance between the earliest and subsequent VP1 variants. As the earliest VP1 variant of the GII.4_2006b, we used a sequence from a May 2006 sample, which was collected in spring about 5 months before the onset of the nationwide epidemics of the GII.4_2006b in October of 2006 in Japan (Motomura et al., 2008).

For the S domain, the Hamming distances of the variants in given seasons were at a constant peak of 0 for 5 years (**Figure 5B**,

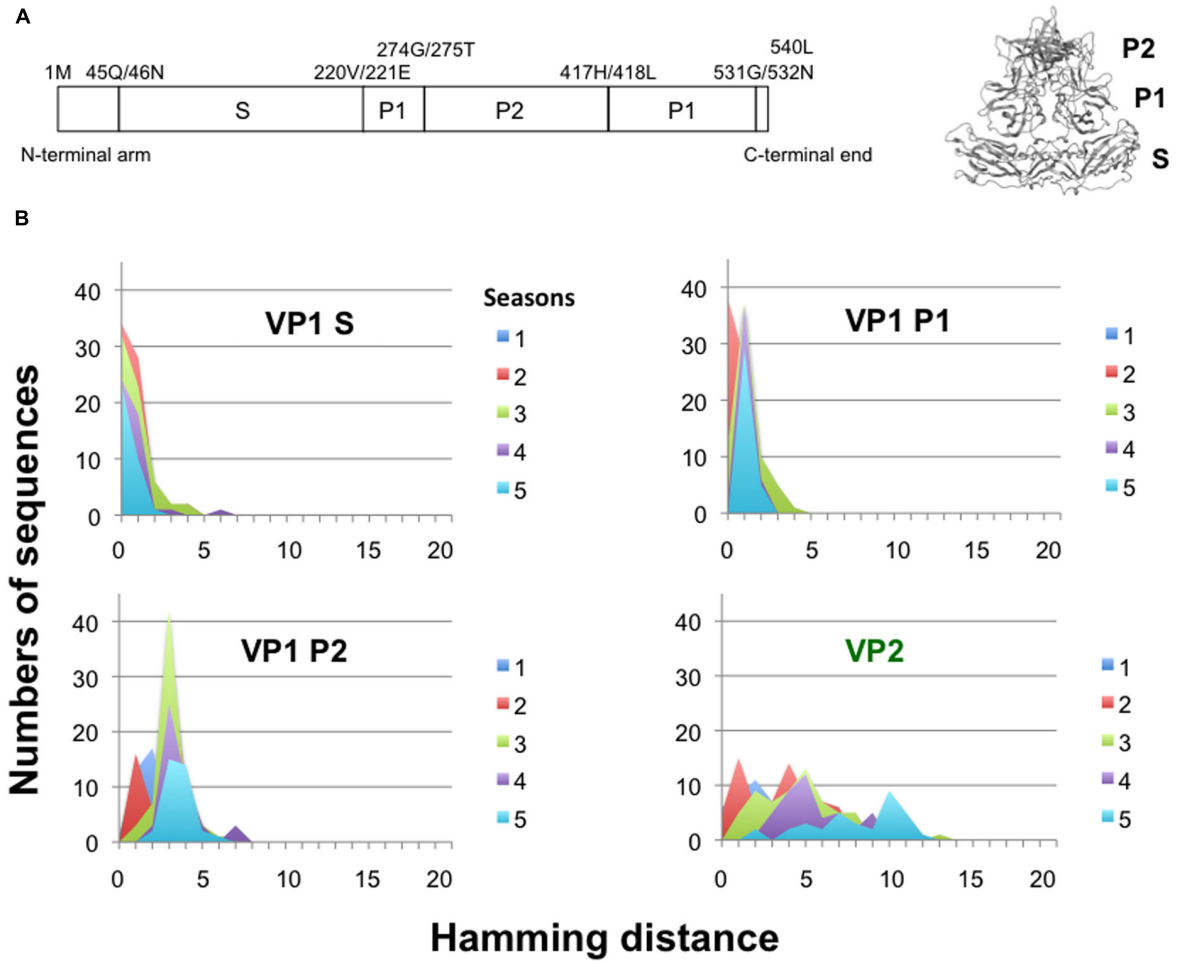


FIGURE 5 | Temporal change in Hamming distance for the NoV GII.4_2006b capsid protein VP1. (A) Schematic illustration of the locations of the S, P1, and P2 domains in the ORF2 (Left) and the VP1 3-D model (Right) of the NoV GII.4_2006b. **(B)** Temporal changes in the Hamming distance. A GII.4_2006b genome from a stool specimen collected in May 2006 in Toyama (accession number AB447443) was the oldest in our GII.4_2006b genome dataset. The sequence was used to calculate the Hamming distance between the earliest and subsequent VP1 variants emerged between October 2006 and March 2011 in Japan. The distributions of the Hamming distances for each season are shown for the S, P1, and P2 domains of the VP1 protein. Temporal changes in the Hamming distance are also shown for the VP2 protein (Lower right).

VP1 Shell). The data suggest that the GII.4_2006b variants having amino acid substitutions in the S domain were mostly cleared during epidemics. For the P1 and P2 subdomains, the peaks of Hamming distances were fixed at 1 and 3 after the second and first epidemic seasons, respectively (**Figure 5B**, VP1 P1 and P2). The data indicate that most of the GII.4_2006b variants in the early epidemics had a few amino acid substitutions in the P domain but they could not accumulate more mutations after the second epidemic season. Thus the P domain was more variable than the S domain in the GII.4_2006b variants, as has generally been documented for other NoVs. However, the accumulation of amino acid substitutions was strictly constrained in the P domain of the GII.4_2006b variants during epidemics. In contrast, the Hamming distances of VP2, a minor structural protein in virion (Glass et al., 2000), continuously increased and showed no evidence of fixation of the peak distance during the study period (**Figure 5B**, VP2).

DISCUSSION

In this report, we studied molecular evolution of the NoV capsid protein of a pandemic lineage, GII.4_2006b. This NoV subpopulation predominated over other coexisting NoV GII.4 subpopulations between the 2006 and 2011 in Japan (**Figure 1**). Notably, the amino acid sequences of variable VP1 protein of the GII.4_2006b populations remained as homogeneous as that of the viral enzymes for the 5 years under an increase in nucleotide diversity (**Figures 2**, **3**). Even the GII.4_2006b population possessing the identical amino acid sequence in the VP1 protein had persisted in the study period (**Figure 4**). Even the hypervariable antigenic P2 subdomain of the VP1 protein had resisted sequential accumulations of amino acid substitutions (**Figure 5**). These results suggest the presence of strong evolutionary constraints on the VP1 protein of the NoV pandemic strain. The finding has implications for

our understanding of molecular evolution, mechanisms of the periodic shifts in the pandemic NoV GII.4 strains, and control of the NoV GII.4 pandemic strain.

First, the finding has implications for understanding fitness landscape and evolution of the VP1 protein of NoV GII.4 pandemic strain. The strong constraints on changes imply that the VP1 protein and enzymes of the GII.4_2006b variants had already gained near maximal levels of fitness benefits in humans around the onset of the outbreaks and that new mutations in the VP1 protein were mostly cleared from the GII.4_2006b population, probably due to a reduction in the viral fitness for the spread in humans. In order to predominate over other coexisting GII.4 variants, the pandemic variant should have the VP1 structure that confers the best ability to evade preexisting herd immunity against NoV at that time, while also having affinity to bind to receptor(s) on human cells. Because the antigenic sites are located near the receptor-binding site, new antigenic mutations always have the risk to attenuate VP1 protein function and thereby to cause reduction in the viral replication fitness in humans. Thus, it is possible that the VP1 protein of the pandemic strain had remained conserved in human populations primarily by the necessity to maintain advantageous physical property of the VP1 protein for immune evasion and infectivity simultaneously.

Secondary, the finding has implications in the periodic shifts of the pandemic NoV GII.4 strains. Provided that the VP1 protein sequence of a given pandemic variant remained conserved following the onset of epidemics as seen in the GII.4_2006b, the human herd immunity against the VP1 protein would become increasingly more effective in association with the spread of the virus in humans. Consequently, niche for the pandemic variant in humans would be reduced, and the pandemic variant eventually be replaced by an alternative variant that has the fittest capsid structure under human herd immunity at that time. Consistently, the numbers of reported NoV infection cases in Japan had decreased annually since the late 2007, and the GII.4_2006b was replaced by a new global pandemic strain GII.4_Sydney 2012 in the 2013/2014 season, as reported in other countries (van Beek et al., 2013; Eden et al., 2014).

Finally, the finding has implications in the control of NoV pandemic strains. Although development of vaccines and antiviral agents are of special importance to reduce damages from the NoV infections, structural variations in the viral proteins can be problematic. In this regard, the present study suggests the presence of strong constraints on changes in capsid protein and enzymes of a NoV GII.4 pandemic variant on the course of 5-year persistence across Japan. The finding provides a rationale for developing vaccines and antiviral agents against a pandemic strain. A basic premise of the control is that the sequences of the VP1 protein and viral enzymes of a given pandemic variant remain highly homogeneous after the onset of pandemic. Therefore, it is important to further accumulate information on the evolution of newly emerged pandemic strains to clarify whether present observations of the amino acid conservation in the VP1 and viral enzymes can be extended to other GII.4 pandemic variants. In parallel, it would be important to study genetic diversity of NoV in nature in order to develop systems to predict a new pandemic variant in advance.

AUTHOR CONTRIBUTIONS

HS conceived the study. MY prepared the computing environment for information science. KK, NT, MN, and TT organized collection of stool specimen. TO, HN, and KM performed sequencing. HN and HS performed variation analysis. HS prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by a Grant-in-Aid for Scientific Research on Innovative Areas to HS (Grant Number: 25115519) from the Japan Society for the Promotion of Science, a Grant-in-Aid for the Research Program on Emerging and Re-emerging Infectious Diseases to HS (Grant Number: 11350406 and 10103800) from the Ministry of Health, Labor and Welfare of Japan, and a Grant-in Aid for the Research Program on Re-emerging Infectious Diseases to MY from the Japan Agency for Medical Research and development, AMED.

ACKNOWLEDGMENTS

We would like to thank the following 20 regional public health institutes for the help of collection of stool specimen: Hokkaido, Aomori Prefecture, Akita Prefecture, Iwate Prefecture, Miyagi Prefecture, Niigata Prefecture, Toyama Prefecture, Fukui Prefecture, Nagano Prefecture, Chiba Prefecture, Aichi Prefecture, Sakai City, Osaka City, Shimane Prefecture, Hiroshima Prefecture, Hiroshima City, Ehime Prefecture, Saga Prefecture, Kumamoto City, and Miyazaki Prefecture.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2017.00410/full#supplementary-material

REFERENCES

Bok, K., Abente, E. J., Realpe-Quintero, M., Mitra, T., Sosnovtsev, S. V., Kapikian, A. Z., et al. (2009). Evolutionary dynamics of GII.4 noroviruses over a 34-year period. *J. Virol.* 83, 11890–11901. doi: 10.1128/JVI.00864-09

Bull, R. A., Tu, E. T., Mciver, C. J., Rawlinson, W. D., and White, P. A. (2006). Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. *J. Clin. Microbiol.* 44, 327–333. doi: 10.1128/JCM.44.2.327-333.2006

Chen, R., Neill, J. D., Estes, M. K., and Prasad, B. V. (2006). X-ray structure of a native calicivirus: structural insights into antigenic diversity and host specificity.

Proc. Natl. Acad. Sci. U.S.A. 103, 8048–8053. doi: 10.1073/pnas.0600421103

Chen, R., Neill, J. D., Noel, J. S., Hutson, A. M., Glass, R. I., Estes, M. K., et al. (2004). Inter- and intragenus structural variations in caliciviruses and their functional implications. *J. Virol.* 78, 6469–6479. doi: 10.1128/JVI.78.12.6469-6479.2004

Debbink, K., Donaldson, E. F., Lindesmith, L. C., and Baric, R. S. (2012). Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity. *J. Virol.* 86, 1214–1226. doi: 10.1128/JVI.06189-11

Donaldson, E. F., Lindesmith, L. C., Lobue, A. D., and Baric, R. S. (2010). Viral shape-shifting: norovirus evasion of the human immune system. *Nat. Rev. Microbiol.* 8, 231–241. doi: 10.1038/nrmicro2296

Eden, J. S., Hewitt, J., Lim, K. L., Boni, M. F., Merif, J., Greening, G., et al. (2014). The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. *Virology* 450–451, 106–113. doi: 10.1016/j.virol.2013.12.005

Eden, J. S., Tanaka, M. M., Boni, M. F., Rawlinson, W. D., and White, P. A. (2013). Recombination within the pandemic norovirus GII.4 lineage. *J. Virol.* 87, 6270–6282. doi: 10.1128/JVI.03464-12

Gallimore, C. I., Iturriza-Gomara, M., Xerry, J., Adigwe, J., and Gray, J. J. (2007). Inter-seasonal diversity of norovirus genotypes: emergence and selection of virus variants. *Arch. Virol.* 152, 1295–1303. doi: 10.1007/s00705-007-0954-9

Glass, P. J., White, L. J., Ball, J. M., Leparc-Goffart, I., Hardy, M. E., and Estes, M. K. (2000). Norwalk virus open reading frame 3 encodes a minor structural protein. *J. Virol.* 74, 6581–6591. doi: 10.1128/JVI.74.14.6581-6591.2000

Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537, 39–64. doi: 10.1007/978-1-59745-251-9_3

Lindesmith, L. C., Beltramello, M., Donaldson, E. F., Corti, D., Swanstrom, J., Debbink, K., et al. (2012a). Immunogenetic mechanisms driving norovirus GII.4 antigenic variation. *PLoS Pathog.* 8:e1002705. doi: 10.1371/journal.ppat.1002705

Lindesmith, L. C., Costantini, V., Swanstrom, J., Debbink, K., Donaldson, E. F., Vinje, J., et al. (2013). Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes. *J. Virol.* 87, 2803–2813. doi: 10.1128/JVI.03106-12

Lindesmith, L. C., Debbink, K., Swanstrom, J., Vinje, J., Costantini, V., Baric, R. S., et al. (2012b). Monoclonal antibody-based antigenic mapping of norovirus GII.4-2002. *J. Virol.* 86, 873–883. doi: 10.1128/JVI.06200-11

Lindesmith, L. C., Donaldson, E. F., and Baric, R. S. (2011). Norovirus GII.4 strain antigenic variation. *J. Virol.* 85, 231–242. doi: 10.1128/JVI.01364-10

Lindesmith, L. C., Donaldson, E. F., Lobue, A. D., Cannon, J. L., Zheng, D. P., Vinje, J., et al. (2008). Mechanisms of GII.4 norovirus persistence in human populations. *PLoS Med.* 5:e31. doi: 10.1371/journal.pmed.0050031

Lopman, B., Vennema, H., Kohli, E., Pothier, P., Sanchez, A., Negredo, A., et al. (2004). Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. *Lancet* 363, 682–688. doi: 10.1016/S0140-6736(04)15641-9

Motomura, K., Oka, T., Yokoyama, M., Nakamura, H., Mori, H., Ode, H., et al. (2008). Identification of monomorphic and divergent haplotypes in the 2006-2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history. *J. Virol.* 82, 11247–11262. doi: 10.1128/JVI.00897-08

Motomura, K., Yokoyama, M., Ode, H., Nakamura, H., Mori, H., Kanda, T., et al. (2010). Divergent evolution of norovirus GII/4 by genome recombination from May 2006 to February 2009 in Japan. *J. Virol.* 84, 8085–8097. doi: 10.1128/JVI.02125-09

Naganawa, S., Yokoyama, M., Shiino, T., Suzuki, T., Ishigatsubo, Y., Ueda, A., et al. (2008). Net positive charge of HIV-1 CRF01_AE V3 sequence regulates viral sensitivity to humoral immunity. *PLoS ONE* 3:e3206. doi: 10.1371/journal.pone.0003206

Noel, J. S., Fankhauser, R. L., Ando, T., Monroe, S. S., and Glass, R. I. (1999). Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. *J. Infect. Dis.* 179, 1334–1344. doi: 10.1086/314783

Oka, T., Yokoyama, M., Katayama, K., Tsunemitsu, H., Yamamoto, M., Miyashita, K., et al. (2009). Structural and biological constraints on diversity of regions immediately upstream of cleavage sites in calicivirus precursor proteins. *Virology* 394, 119–129. doi: 10.1016/j.virol.2009.08.018

Prasad, B. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., and Estes, M. K. (1999). X-ray crystallographic structure of the Norwalk virus capsid. *Science* 286, 287–290. doi: 10.1126/science.286.5438.287

Prasad, B. V., Rothnagel, R., Jiang, X., and Estes, M. K. (1994). Three-dimensional structure of baculovirus-expressed Norwalk virus capsids. *J. Virol.* 68, 5117–5125.

Reuter, G., Krisztalovics, K., Vennema, H., Koopmans, M., and Szucs, G. (2005). Evidence of the etiological predominance of norovirus in gastroenteritis outbreaks–emerging new-variant and recombinant noroviruses in Hungary. *J. Med. Virol.* 76, 598–607. doi: 10.1002/jmv.20403

Robilotti, E., Deresinski, S., and Pinsky, B. A. (2015). Norovirus. *Clin. Microbiol. Rev.* 28, 134–164. doi: 10.1128/CMR.00075-14

Shannon, C. E. (1997). The mathematical theory of communication. 1963. *MD Comput.* 14, 306–317.

Siebenga, J. J., Vennema, H., Renckens, B., De Bruin, E., Van Der Veer, B., Siezen, R. J., et al. (2007). Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. *J. Virol.* 81, 9932–9941. doi: 10.1128/JVI.00674-07

Takahata, T., Takeda, E., Tobiume, M., Tokunaga, K., Yokoyama, M., Huang, Y. L., et al. (2017). Critical contribution of Tyr15 in the HIV-1 integrase (IN) in facilitating IN assembly and nonenzymatic function through the IN precursor form with reverse transcriptase. *J. Virol.* 91:e02003-16. doi: 10.1128/jvi.02003-16

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

van Beek, J., Ambert-Balay, K., Botteldoorn, N., Eden, J. S., Fonager, J., Hewitt, J., et al. (2013). Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012. *Euro Surveill.* 18, 8–9.

Vipond, I. B., Caul, E. O., Hirst, D., Carmen, B., Curry, A., Lopman, B. A., et al. (2004). National epidemic of Lordsdale Norovirus in the UK. *J. Clin. Virol.* 30, 243–247. doi: 10.1016/j.jcv.2003.11.011

Zheng, D. P., Ando, T., Fankhauser, R. L., Beard, R. S., Glass, R. I., and Monroe, S. S. (2006). Norovirus classification and proposed strain nomenclature. *Virology* 346, 312–323. doi: 10.1016/j.virol.2005.11.015

Zheng, D. P., Widdowson, M. A., Glass, R. I., and Vinje, J. (2010). Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006. *J. Clin. Microbiol.* 48, 168–177. doi: 10.1128/JCM.01622-09

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Sato, Yokoyama, Nakamura, Oka, Katayama, Takeda, Noda, Tanaka and Motomura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

METHODS
published: 26 August 2016
doi: 10.3389/fmicb.2016.01294

Whole Genome Sequencing of *Enterovirus species C* Isolates by High-Throughput Sequencing: Development of Generic Primers

Maël Bessaud[1,2,3]*, *Serge A. Sadeuh-Mba*[4], *Marie-Line Joffret*[1,2,3], *Richter Razafindratsimandresey*[5], *Patsy Polston*[1,2], *Romain Volle*[1,2], *Mala Rakoto-Andrianarivelo*[6], *Bruno Blondel*[1,2], *Richard Njouom*[4] and *Francis Delpeyroux*[1,2,3]

[1] Unité de Biologie des Virus Entériques, Institut Pasteur, Paris, France, [2] Institut National de la Santé et de la Recherche Médicale, U994, Paris, France, [3] WHO Collaborating Center for Research on Enteroviruses and Viral Vaccines, Institut Pasteur, Paris, France, [4] Centre Pasteur du Cameroun, Service de Virologie, Yaoundé, Cameroon, [5] Unité de virologie, Institut Pasteur de Madagascar, Antananarivo, Madagascar, [6] Centre d'infectiologie Charles-Mérieux, Université Ankatso, Antananarivo, Madagascar

OPEN ACCESS

Edited by:
Akio Adachi,
University of Tokushima, Japan

Reviewed by:
Daniel C. Pevear,
VenatoRx Pharmaceuticals, USA
Chuan Xiao,
University of Texas at El Paso, USA
Sindy Böttcher,
Robert Koch Institute, Germany

***Correspondence:**
Maël Bessaud
mael.bessaud@pasteur.fr

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *26 May 2016*
Accepted: *05 August 2016*
Published: *26 August 2016*

Citation:
Bessaud M, Sadeuh-Mba SA, Joffret M-L, Razafindratsimandresey R, Polston P, Volle R, Rakoto-Andrianarivelo M, Blondel B, Njouom R and Delpeyroux F (2016) Whole Genome Sequencing of Enterovirus species C Isolates by High-Throughput Sequencing: Development of Generic Primers. Front. Microbiol. 7:1294. doi: 10.3389/fmicb.2016.01294

Enteroviruses are among the most common viruses infecting humans and can cause diverse clinical syndromes ranging from minor febrile illness to severe and potentially fatal diseases. *Enterovirus species C* (EV-C) consists of more than 20 types, among which the three serotypes of polioviruses, the etiological agents of poliomyelitis, are included. Biodiversity and evolution of EV-C genomes are shaped by frequent recombination events. Therefore, identification and characterization of circulating EV-C strains require the sequencing of different genomic regions. A simple method was developed to quickly sequence the entire genome of EV-C isolates. Four overlapping fragments were produced separately by RT-PCR performed with generic primers. The four amplicons were then pooled and purified prior to being sequenced by a high-throughput technique. The method was assessed on a panel of EV-Cs belonging to a wide-range of types. It can be used to determine full-length genome sequences through *de novo* assembly of thousands of reads. It was also able to discriminate reads from closely related viruses in mixtures. By decreasing the workload compared to classical Sanger-based techniques, this method will serve as a precious tool for sequencing large panels of EV-Cs isolated in cell cultures during environmental surveillance or from patients, including vaccine-derived polioviruses.

Keywords: *Enterovirus species C*, poliovirus, high-throughput sequencing, recombination, *de novo* assembly

INTRODUCTION

The members of the *Enterovirus species C* (EV-C), genus *Enterovirus*, family *Picornaviridae*, are non-enveloped viruses with a single positive strand RNA genome. The virions contain one copy of the genome, which is about 7500 nucleotides in length and consists of two untranslated regions (5′- and 3′-UTR) flanking a unique large open reading frame. The polyprotein encoded by this open reading frame is first cleaved into three precursors (P1–P3) that are subsequently cleaved into

functional proteins. P1 gives rise to the four structural capsid proteins (VP1–VP4) while P2 and P3 generate non-structural proteins involved in the viral cycle. Enteroviruses cause a wide spectrum of human diseases, with clinical signs ranging from mild febrile illness, such as the common cold, to severe forms, such as acute haemorrhagic conjunctivitis, myocarditis, encephalitis, and acute flaccid paralysis (Tapparel et al., 2013).

Currently, more than 20 types of EV-Cs have been identified (Knowles et al., 2012), including the three serotypes of poliovirus (PV-1 to −3) that can induce severe and potentially fatal cases of poliomyelitis in humans. Typing of enteroviruses relies on molecular characterization of the capsid-encoding region (Blomqvist and Roivainen, 2016). For this purpose, many generic assays were developed to amplify genomic fragments within the VP4-, VP2-, or VP1-encoding regions by RT-PCR and to sequence the resulting amplicons. The sequences of clinical or field viruses can be compared to those of prototype strains in order to determine their respective types.

Besides typing, molecular characterization of circulating EV-Cs generally implies sequencing of untranslated or non-structural genomic regions, or even full-length sequencing of the genome. Indeed, recombination events between enteroviruses are known to be very frequent, thus leading to complex ecosystems of co-circulating viruses featuring mosaic genomes (Savolainen-Kopra and Blomqvist, 2010; Combelas et al., 2011; Kyriakopoulou et al., 2015). As recombination constitutes a powerful force that drives enterovirus evolution, sequencing of different genomic regions of isolates is crucial to detect recombination events. Such events are revealed by incongruent clustering of genetic sequences in phylogenetic trees, depending on the studied genomic regions.

Generic assays have been developed to amplify non-structural regions of EV genomes, particularly the 5′-UTR and the 2C- and the 3D-encoding regions (Bessaud et al., 2008). These wide-range assays can be used to characterize portions of the genome of EV-Cs but determining the full-length genomic sequences requires additional sequencing steps to bridge the gaps between the sequences determined through generic assays. These steps require the usage of specific primers that have to be designed for each isolate and can be very labor-intensive when numerous isolates of many types are studied.

In order to facilitate the full-length sequencing of EV-C isolates, a simple method was developed, which allowed the amplification of the whole genome by RT-PCR using four generic primer pairs. Four overlapping fragments were produced by performing four separate PCR reactions on cDNAs generated through a single RT reaction. The four amplicons were then pooled and purified prior to sequencing by high-throughput sequencing methods.

The amplification step was assessed on a panel of prototype and field strains whose genome had been previously sequenced by the Sanger method. Sequencing data was analyzed by mapping the reads against reference sequences and by *de novo* assembly. This method was able to determine the full-length genomic sequences of EV-C isolates with high sensitivity and to discriminate the different genomic sequences within mixtures of viruses.

TABLE 1 | Viruses used in the study.

Type	Strain or isolate	Origin	Cycle threshold[a]
CV-A1	Tompkins	Prototype strain	14.5 ± 0.2
CV-A11	Belgium	Prototype strain	16.6 ± 0.2
CV-A11	G9	Prototype strain	14.9 ± 0.7
CV-A11	66122	From Rakoto-Andrianarivelo et al. (2007)	15.6 ± 0.1
CV-A11	66990	From Rakoto-Andrianarivelo et al. (2007)	15.9 ± 0.1
CV-A13	Flores	Prototype strain	15.7 ± 0.3
CV-A13	G13	Prototype strain	23.7 ± 0.1
CV-A13	67900	From Rakoto-Andrianarivelo et al. (2007)	17.1 ± 0.4
CV-A13	67001	From Rakoto-Andrianarivelo et al. (2007)	13.3 ± 0.2
CV-A17	G12	Prototype strain	13.8 ± 0.3
CV-A17	67591	From Rakoto-Andrianarivelo et al. (2007)	14.4 ± 0.9
CV-A17	68154	From Rakoto-Andrianarivelo et al. (2007)	15.2 ± 0.2
CV-A19	NIH-8663	Prototype strain	23.2 ± 1.0
CV-A20	Cecil	Prototype strain	16.0 ± 0.1
CV-A20	IH35	Prototype strain	13.5 ± 0.4
CV-A20	Tulane	Prototype strain	16.1 ± 0.1
CV-A21	Coe	Prototype strain	17.7 ± 0.3
EV-C95	T08-083	From Sadeuh-Mba et al. (2013)	14.5 ± 0.5
EV-C99	68229	From Rakoto-Andrianarivelo et al. (2007)	11.5 ± 0.1
PV-1	Sabin	Vaccine strain	15.6 ± 0.1
PV-2	Sabin	Vaccine strain	17.4 ± 0.3
PV-3	Sabin	Vaccine strain	20.9 ± 0.1

[a] *Cycle threshold values obtained by a pan-enterovirus real-time RT-PCR assay performed on the extracted RNAs.*

MATERIALS AND METHODS

Viruses

Twenty-two viruses were used for this study (**Table 1**). The prototype viruses are available on the Pasteur Institute collection (*Centre de ressources biologiques de l'Institut Pasteur*, Paris, France). Field isolates originated from stool samples collected in Madagascar in 2002 (Rakoto-Andrianarivelo et al., 2005, 2007) and in Chad in 2008–2009 (Sadeuh-Mba et al., 2013). The whole genomic sequence of these isolates were previously determined by the Sanger method (Bessaud et al., 2011).

All work with infectious viruses was carried out in a BSL-2 facility. All viruses were grown in HEp-2c cell monolayers in DMEM supplemented with 2% fetal calf serum and 2 mM L-glutamine at 37°C, except coxsackievirus (CV) A1 Tompkins and CV-A19 NIH-8663, see below. These latter cannot be propagated in cell lines but can infect suckling mice. RNA of these two viruses was extracted from brains of mice inoculated intracerebrally. These brains were retrieved from the laboratory collection and no mice were used during this study. Brains were crushed in PBS before RNA extraction.

Sequence Analysis and Primer Design

EV-C full-length nucleotide sequences retrieved from the GenBank database were aligned using CLC Main Workbench 7.6.4 software (CLC bio). Eight degenerated primers were designed to target conserved genomic regions (**Table 2**).

TABLE 2 | Primers used in this study.

Step	Name	5′ →3′ sequence	Genome position[a]
cDNA synthesis	heptaN	NNNNNNN	
PCR	C004	TTAAAACAGCYYKDGGGTTG	1–20
	C005	CCGAATYAAARRAAAATTTACCC	7437–7415
	C008	CARTTYAAGASCAARCAYCG	4459–4478
	C009	ACCATYTGRCARAANARYTTCA	4703–4682
	C018	ATGYTNGGNACNCAYNTNATHTGGGA	2212–2237
	C019	CCYTGYTCCATNRCYTCHTCYTC	3833–3811
	C021	TCDGGNARYTTCCACCACCA	1205–1186
	C022	GARGCNTGYGGDTAYAGYGA	967–986

[a]*Relative to PV-2 strain Sabin.*

RNA Extraction

Viral RNA was extracted from 250 μL of culture supernatants or clarified brain extracts using the High Pure Viral RNA kit (Roche Diagnostics, Meylan, France), following the manufacturer's instructions.

In order to check the RNA extraction step, a one-step real-time RT-PCR assay was performed on extraction products using a pan-enterovirus generic assay previously described (Monpoeho et al., 2000).

Synthesis of the Four Overlapping Amplicons by 2-Step RT-PCR

For each virus, four overlapping DNA fragments were produced by RT-PCR (**Figure 1**). cDNA synthesis was performed as previously described (Bessaud et al., 2008). The reaction mixture contained 5 μL of purified viral RNA, 2 μL of 5X First-Strand Buffer, 0.01 M dithiothreitol (1 μL), 100 ng of the random primers heptaN (1 μL), 10 nmol of each dNTP (1 μL of a 10 mM mixture), and 100 U of SuperScript II (0.5 μL). The RT reaction mixture was incubated at 25°C for 10 min, 42°C for 45 min and 95°C for 5 min.

The cDNA was then used as a template for amplification in four PCRs carried out in a final volume of 50 μL that included 5 μL of 10X PCR Buffer w/o $MgCl_2$, 1.5 mM of $MgCl_2$, 10 nmol of each dNTP, 50 pmol of each primer, 2 μL of cDNA, and 2.5 U of Platinum *Taq* DNA polymerase (Invitrogen). The thermocycler profile was 2 min at 94°C followed by 30 cycles of 30 s at 94°C, 30 s at 55°C, and 3 min at 72°C.

Ten microliters of each PCR product were analyzed on ethidium bromide-stained agarose gels. For each virus, the four PCR products were pooled, purified on silica columns (Wizard SV Gel and PCR Clean-Up System, Promega) and eluted in 30 μL of water. No purification of the amplicons was performed by gel-excision, even when additional bands were detected on ethidium bromide-stained agarose gels.

Assay Sensitivity

The sensitivity of the assay was evaluated using serial three-fold dilutions of a cell culture supernatant infected by CV-A13 isolate 67001 whose titer was determined according to the WHO standard protocol (Anonymous, 2004). To maintain the same amount of cellular nucleic acids across dilutions, dilutions were prepared using a supernatant of confluent non-infected HEp-2c cell monolayer that was frozen and thawed twice and then clarified by centrifugation. RNA was extracted from 250 μL of each dilution and subjected to real-time RT-PCR as described in Section RNA Extraction, and amplified by RT and PCR as described in Section Synthesis of the Four Overlapping Amplicons by 2-Step RT-PCR.

Detection of Virus Mixtures

Four mixtures were prepared using cell culture supernatants of CV-A13 isolates 67900 and 67001 whose titer was determined according to the WHO standard protocol (Anonymous, 2004). Titers were adjusted to $10^{7.5}$ $TCID_{50}.mL^{-1}$ before mixing. Different mixtures were prepared with 67900/67001 titer ratios ranging from 1:1 to 10:1. RNA was extracted from 250 μL of these mixtures and subjected to RT and PCR as described in Section Synthesis of the Four Overlapping Amplicons by 2-Step RT-PCR.

Sequencing Process

DNA concentration of the purified RT-PCR products was determined by using a VarioskanLux (ThermoScientific). Libraries were built using 1 ng of DNA with the Nextera XT DNA Library Preparation kit in a SureCycler 8800 thermocycler (Agilent). After purification on AMPure beads (Beckman), the libraries were controlled using the High Sensitivity D1000 assay (Agilent) on a TapeStation 2200. Sizing was achieved by electrophoresis on a PippinPrep System with the PippinPrep kit CDF1510 (Ozyme). Finally, the libraries were quantified using the KAPA Quantification kit on a LightCycler 96 System (Roche). Sequencing was performed on a NextSeq500 with the HighOutput or MidOutput kits. All kits were used following manufacturer's instructions.

Data Analysis

Trimming

Reads were demultiplexed followed by the removal of tags and adaptors. After importation in CLC Genomics Workbench 8.5 (CLCbio), reads were trimmed using the following parameters: Trim quality score limit = 0.01; Trim ambiguous nucleotides with Maximum number of ambiguities = 1. To avoid the presence of primer sequences in the reads, the 26 5′- and 3′-terminal nucleotides were removed. After trimming, reads shorter than 50 nucleotides were discarded.

Mapping of the Reads to a Reference

For each sample, the trimmed reads were mapped against the reference sequence of the corresponding virus. The sequences used as references were the sequences previously determined by the Sanger method. The entire genome of all field isolates were previously sequenced in our laboratory (Bessaud et al., 2011); for prototype strains, the sequences used as references were retrieved from the Genbank database. Mapping was achieved with CLC Genomics Workbench 8.5 by using the following parameters: Mismatch cost = 10; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.5; Similarity fraction = 0.95.

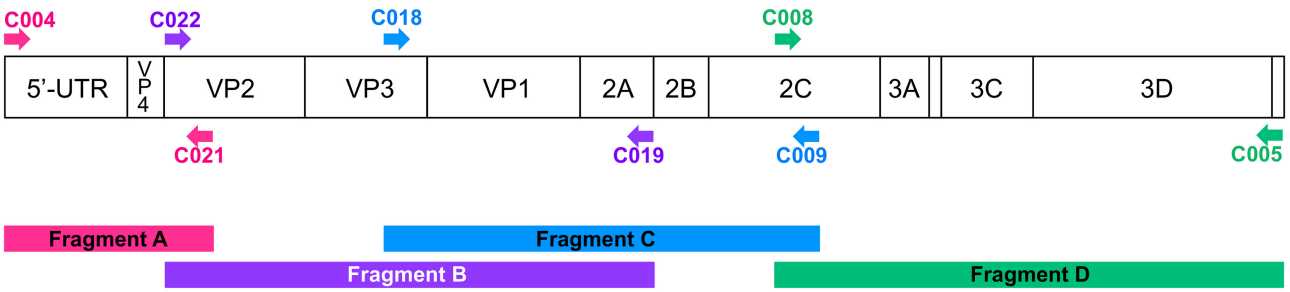


FIGURE 1 | Schematic representation of the locations of the primers used for the PCR step. The upper scheme represents the organization of the enterovirus genomic RNA. Arrows indicate the sites targeted by the primers. The RT-PCR products are colored according to the corresponding primer pair.

De novo Assembly

For each sample, the trimmed reads were assembled without any reference using CLC Genomics Workbench 8.5 with the following parameters: Mismatch cost = 2; Insertion cost = 2; Deletion cost = 2; Length fraction = 0.5; Similarity fraction = 0.95.

All contigs longer than 200 nucleotides were submitted to BLAST analysis (Mount, 2007). Virus contigs were compared with the reference sequence of the corresponding virus determined by the Sanger method through alignment.

RESULTS

Sequence Analysis and Primer Design

Eight degenerated primers were designed to target nucleotide sequences that were conserved among EV-Cs (**Figure 1**). C004 and C005 target the extremity of the 5′ and 3′ non-coding region, respectively. C018, C019, and C008 match with conserved regions already targeted by pan-EV molecular assays (Caro et al., 2001; Nix et al., 2006; Bessaud et al., 2008) within the VP3 gene, at the 2A/2B junction, and within the 2C *cis*-acting replication element (Cordey et al., 2008), respectively. The three other primers, C021, C022, and C009 match with VP2 and 2C sequences.

The primer pairs C004/C021, C022/C019, C018/C009, and C008/C005 were used to produce four overlapping DNA fragments (fragments A–D) that span the entire viral genome. Fragments ranged from ∼1200 to ∼3000 nucleotides in length. Overlapping regions were approximately 240-nt-long for fragments A and B and fragments C and D (**Figure 1**). The overlapping region for fragments B and C was longer than 1600 nt.

RT-PCR Amplification

The four primer pairs C004/C021, C022/C019, C018/C009, and C008/C005 were tested on prototype and field strains representative of 12 EV-C types (**Table 1**). This panel included the three PV serotypes and non-polio types commonly isolated during environmental or epidemiological surveillance, such as CV-A11, CV-A13, CV-A17, CV-A20, CV-A21, and EV-C99. One field strain belonged to the type EV-C95 that was recently discovered (Junttila et al., 2015) and of which few isolates were reported in Africa. The panel also included the prototype strains of CV-A1 and CV-A19, two types that are uncultivable on cell lines (Brown et al., 2003). Overall, 22 viruses were used to assess the efficiency of the primers. After RNA extraction, detection of viral RNA by real-time RT-PCR resulted in cycle threshold values ranging from 11.5 ± 0.1 to 23.7 ± 0.1 (**Table 1**).

Following the RT step with random primers performed on viral RNA extracted from infected cell culture supernatants (or from crushed brains of mice infected by CV-A1 or CV-A19), the PCR performed with the four primer pairs produced gel bands at the expected size for all the viruses (data not shown). Some additional bands were also observed for certain RT-PCR products. Among our panel, CV-A19 NIH-8663 RNA gave the weakest bands on ethidium bromide-stained gel after the RT-PCR step. It is not possible to exclude the hypothesis that the four primer pairs had a low efficiency in amplifying the genome of this virus through PCR because of mismatches. Nonetheless, since the RNA of this virus was extracted from a mouse brain collected more than 20 years ago, this result is more likely due to the low amount of full-length virus RNA in this sample.

Results of Sequencing

The products of RT-PCR amplification of non-polio EV were sequenced by Illumina technology. Due to French regulations relating to PV containment, the sequencing platform was not allowed to handle the PV RT-PCR products.

For each sample, the four RT-PCR fragments were produced separately and then pooled prior to purification. Sequencing resulted in a list of reads of which most were ∼100 nucleotide-long after trimming. Because all of the samples were not sequenced during the same run, the number of reads varied from sample to sample. Thus, the number of reads after trimming ranged from 76,802 to 714,322 (**Table 3**).

Bioinformatics Analysis

Raw data generated by high-throughput sequencing was analyzed by two different methods commonly used for virus characterization.

The first method consists in mapping the reads against a reference sequence. This method can be used to re-sequence isolates or to detect single-point mutations that appear when a given virus is propagated under different conditions (e.g., various

TABLE 3 | Overview of the results of the bioinformatics analysis performed on sequencing data.

Strain or isolate	Total number of reads	Mapping on reference sequence		*De novo* assembly	
		Number of mapped reads (%[a])	Average depth	Number of reads in the virus contig (%[b])	Number of host contigs
CV-A1					
Tompkins	391,846	380,207 (97%)	7032	379,991 (96%)	200
CV-A11					
Belgium	373,748	337,637 (90%)	6214	337,708 (90%)	232
G9	121,080	119,260 (98%)	2310	120,181 (99%)	16
66122	76,802	76,342 (99%)	982	76,290 (99%)	3
66990	121,784	114,961 (94%)	2277	120,712 (99%)	10
CV-A13					
Flores	88,094	86,883 (97%)	1608	81,912 (93%)	6
G13	101,372	98,056 (96%)	1902	100,746 (99%)	12
67900	175,426	174,488 (99%)	3382	154,526 (88%)	1
67001	189,890	185,674 (97%)	3689	188,303 (99%)	0
CV-A17					
G12	554,763	553,436 (99%)	10,404	553,221 (99%)	28
67591	421,668	367,691 (87%)	4741	365,019 (87%)	197
68154	84,786	83,710 (98%)	1548	83,251 (98%)	14
CV-A19					
NIH-8663	374,976	153,262 (40%)	2958	152,980 (40%)	783
CV-A20					
Cecil	619,194	616,252 (99%)	11,525	615,858 (99%)	59
IH35	698,498	630,711 (99%)	11,657	632,064 (99%)	39
Tulane	714,322	712,310 (99%)	13,336	712,445 (99%)	21
CV-A21					
Coe	570,328	525,606 (92%)	9608	565,659 (99%)	115
EV-C95					
T08-083	96,460	96,059 (99%)	1776	96,110 (99%)	7
EV-C99					
68229	575,410	573,084 (99%)	10,435	572,958 (99%)	19

[a] *These percentages indicate the proportion of total reads that matched with the reference sequence.*
[b] *These percentages indicate the proportion of total reads that were included in the virus contig through de novo assembly.*

host species, various cell lines or presence of antiviral molecules, for example).

The second method consists in constructing contigs through *de novo* assembly (i.e., without any reference sequence to drive the assembly). This method can be used to determine the full-length sequence of viruses isolated from clinical or environmental samples.

Mapping on Reference Sequence

For each sample, reads were mapped against the corresponding sequence previously determined by the Sanger method (**Table 3**). For all samples except one, more than 87% of the reads mapped against the reference sequence. The only exception was CV-A19 NIH-8663, a non-cultivable virus, in which only 40% of the reads mapped to the reference.

For all samples, the reference sequence was fully covered without any gaps (breadth of coverage >99.0%). The breadth of coverage never reached 100.0% because the sequences corresponding to the outer primers C004 and C005 and the few nucleotides located downstream C005 were absent in the final contig.

Average depth of coverage, which is the average number of times a base is sequenced, was very high (>980) but varied greatly from sample to sample (**Table 3**), depending on the total number of reads (correlation coefficient >0.999). Coverage depth was heterogeneous along the mapping (**Figure 2**) but was high enough to deduce an unambiguous consensus sequence even for samples with a low number of reads (CV-A11 66122 or CV-A17 68154 for instance) or a low proportion of mapping reads (CV-A19 NIH-8663).

The consensus sequences of field strains deduced from mapping were compared to the consensus sequences determined by the Sanger method using RNA extracted from the same cell culture supernatants. For each virus, the two consensus sequences were identical indicating that high-throughput sequencing did not introduced nucleotide

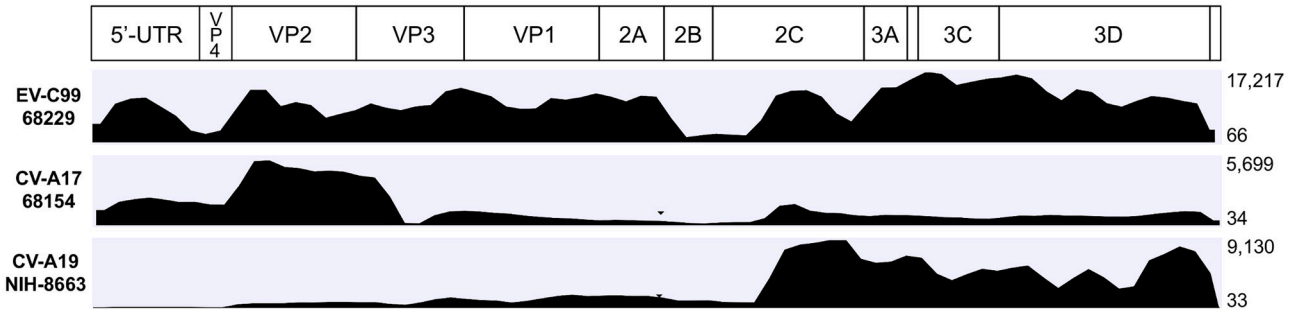


FIGURE 2 | Coverage depth observed for three samples after mapping on their respective reference sequence. Coverage depth is the number of times a base is sequenced. On the right side of each graph, the minimum and maximum coverage depth values are indicated.

changes in the consensus sequences compared to the Sanger method.

De novo Assembly

In order to assess the ability of the method to determine full-length sequences of the original isolates, the trimmed reads of each sample were assembled without using a reference.

For all samples, the *de novo* assembly tool of CLC Genomics Workbench succeeded in assembling the virus genome into a unique contig that covered the entire genome between the outer primers C004 and C005. The number of reads included in the final full-length contigs was virtually similar to the number of reads previously found to map against the reference (**Table 3**). The consensus sequence deduced from *de novo* assembly was identical to the consensus sequence already determined by mapping.

For most samples, besides the viral contigs, *de novo* assembly produced additional contigs that were identified as human sequences (or mouse sequences for CV-A1 and CV-A19 samples) by BLAST analysis. These contigs were from non-specific amplification of cellular nucleic acids during the RT-PCR process or from contamination of the viral RNA with cellular DNA during extraction. The number of such contigs varied tremendously from sample to sample, ranging from 0 to 783 (**Table 3**). Nonetheless, contaminating nucleic acids did not compromise the proper assembly of the viral genome, even for the CV-A19 sample that contained mostly non-viral reads.

Sensitivity of the Assay

In order to evaluate the sensitivity of the assay, three-fold serial dilutions of a CV-A13 67001 supernatant at $10^{7.8}$ $TCID_{50}.mL^{-1}$ were prepared. After RNA extraction, detection of viral RNA by real-time RT-PCR showed positive results for extracts from the first 15 dilutions (**Table 4**).

The 2-step RT-PCR assays using C004/C021, C022/C019, C018/C009, and C008/C005 were tested on these extracts. The four primer pairs were able to generate detectable amplicons for RNA extracted from the undiluted supernatant and from the first four dilutions (**Figure 3**). Since ethidium bromide-stained gels have a detection limit of few nanograms of DNA per band, the absence of visible bands after the amplification of the highest dilutions does not indicate necessarily that no amplicons were generated for these dilutions. Therefore, all the RT-PCR products were investigated through high-throughput sequencing.

After pooling and purification, the RT-PCR products of each dilution were sequenced during the same run. The total number of reads varied from sample to sample, from 88,930 to 316,336 (**Table 4**). In the sample from the undiluted supernatant, 99.9% of reads mapped against the reference sequence. As expected this proportion decreased when the dilution factor increased but even the samples from the highest dilutions contained a few hundred reads from the virus genome. The non-mapped reads were from cellular nucleic acids. For the six last samples (dilution factor $\geq 3^{10}$), the number of virus reads was too low to reconstitute the full-length genome sequence through *de novo* assembly. By contrast, *de novo* assembly succeeded in building the full-length genome for all samples with dilution factor $\leq 3^9$, including the last one in which only 2.5% of the reads were from the virus. For this sample, the coverage depth ranged from 5 to 95 along the contig (average of 28). In spite of this relatively low depth, the contig sequence generated for this sample was identical to the Sanger consensus sequence.

These results demonstrated that the method was able to generate full-length sequences of EV-C isolates from a few thousand reads and indicated that sequencing can be attempted even in the absence of detectable bands on agarose gels.

Detection of Mixtures

Supernatants of cell cultures inoculated with clinical and environmental samples often contain mixtures of enteroviruses. In order to determine whether the method was able to sequence mixtures by assembling properly different viral contigs, four mixtures of viruses were prepared by mixing cell culture supernatants of two viruses, 67900 and 67001. These two viruses belong to the same serotype (CV-A13) and display together a nucleotide identity of 83%. The four mixtures contained the same volume of 67900 supernatant but different volumes of 67001 supernatant (**Table 5**). Thus, the titer ratios ranged from 1:1 to 10:1.

For the four mixtures, *de novo* assembly generated separate contigs for reads from 67900 and from 67001. As expected, the proportion of reads included in the 67900 contigs increased from mixture 1 to mixture 4 while the proportion of reads in the 67001 contigs decreased (**Table 5**). Nonetheless, even in mixture

TABLE 4 | Evaluation of the sensitivity of the assay.

Dilution factor	Viral titer ($TCID_{50}.mL^{-1}$)	RT-PCR cycle threshold	DNA concentration ($ng.\mu L^{-1}$) [b]	Number of reads	Number of reads mapping against the reference (% [c])	Construction of a full-length contig through *de novo* assembly
3^0	$10^{7.8}$	16.2 ± 0.1	>200.0	197,190	196,993 (99.9%)	Yes
3^1	$10^{7.3}$	17.9 ± 0.4	90.4	143,770	143,125 (99.5%)	Yes
3^2	$10^{6.8}$	20.1 ± 0.4	76.7	140,554	138,103 (98.2%)	Yes
3^3	$10^{6.4}$	21.6 ± 0.2	18.7	112,070	104,730 (93.4%)	Yes
3^4	$10^{5.9}$	23.3 ±0.1	6.8	190,346	172,273 (90.5%)	Yes
3^5	$10^{5.4}$	25.1 ± 0.1	5.5	316,336	218,511 (69.1%)	Yes
3^6	$10^{4.9}$	26.7 ± 0.1	3.9	216,594	117,302 (54.1%)	Yes
3^7	$10^{4.5}$	28.4 ± 0.2	3.4	117,920	48,054 (27.0%)	Yes
3^8	$10^{4.0}$	30.0 ± 0.2	3.8	130,334	14,828 (11.3%)	Yes
3^9	$10^{3.5}$	31.6 ± 0.1	2.8	88,930	2282 (2.5%)	Yes
3^{10}	$10^{3.0}$	33.4 ± 0.2	2.5	119,950	1109 (0.9%)	No
3^{11}	$10^{2.6}$	35.2 ± 0.6	3.1	104,470	511 (0.4%)	No
3^{12}	$10^{2.1}$	36.1 ± 0.3	2.7	124,876	297 (0.2%)	No
3^{13}	$10^{1.6}$	37.5 ± 0.6	1.1	105,656	208 (0.2%)	No
3^{14}	$10^{1.1}$	> 38.0[a]	1.5	176,260	314 (0.2%)	No
3^{15}	$10^{0.6}$	> 39.0[a]	1.1	111,890	228 (0.2%)	No

[a] *Only one well of the triplicate assay showed a detectable increase in fluorescence.*
[b] *After pooling of the four RT-PCR products and purification, the DNA concentration was measured on a VarioskanLux spectrophotometer.*
[c] *These percentages indicate the proportion of total reads that matched with the reference sequence.*

4, which contained 10-fold less starting 67001 RNA compared to mixture 1, the method was able to properly generate a full-length contig corresponding to the expected sequence of 67001.

These results demonstrated that, by using the parameters indicated in Section *De novo* Assembly, the stringency of the *de novo* bioinformatics analysis was high enough to allow the segregation of the reads into separate contigs.

DISCUSSION

High-throughput sequencing techniques constitute a powerful tool for the study of viruses (Quiñones-Mateu et al., 2014; Nelson and Hughes, 2015). By allowing concomitant sequencing of millions of DNA fragments, they allow rapid sequencing of a great number of samples and in-depth characterization of minority genomic variants. The aim of this study was to develop a convenient method allowing the whole sequencing of the genome of EV-C isolates by using high-throughput sequencing.

Different strategies have been reported to copy viral genomic RNAs into DNA fragments that can be subsequently sequenced by high-throughput techniques.

Some strategies are based on random amplification using non-specific primers (Berthet et al., 2008; Djikeng et al., 2008). These strategies generally lead to the generation of DNA libraries that mainly consist of non-viral sequences, thus decreasing the amount of relevant reads obtained by sequencing. Reducing the amount of unwanted reads requires the use of additional procedures to physically enrich the samples for viral RNAs (Hall et al., 2014) or to limit the amplification of host nucleic acids (Ge et al., 2015). Whole sequencing of virus RNA genomes amplified by random primers can also be impaired by the relatively low amplification rate of some genomic regions, which can lead to gaps in the genomic sequences (Rosseel et al., 2013).

Alternate strategies are based on primers that specifically target the viral RNAs to be sequenced. Such strategies have been already reported to sequence several positive-strand RNA viruses, including enterovirus A71 (Wright et al., 2011; Baronti et al., 2015; Cruz et al., 2016; Thomson et al., 2016). In our experiments, using EV-C-targeting primers rather than random primers limited the number of reads from cellular nucleic acids: the proportion of reads from viral origin was higher than 90% for most samples. The method was thus sensitive enough to determine full-length genomes through *de novo* assembly from only a few thousand reads.

The four primer pairs were efficiently tested on viruses belonging to EV-C types commonly isolated during epidemiological studies, including the three serotypes of PV. They also amplified CV-A1 and CV-A19, which were more closely related to EV-C types recently identified, such as EV-C113 and EV-C116 (Tokarz et al., 2013). Since the primers were designed from an alignment comprising the sequences of all currently known EV-C types, they are likely to be used successfully to amplify any EV-Cs. We cannot exclude that these primers are also able to amplify genomic sequences of viruses closely related to EV-Cs. In particular, the primer C008 targets the *cis*-acting replication element that is highly conserved among EV-As, -Bs, -Cs, and –Ds (Cordey et al., 2008). Reads from EV-As, -Bs, or –Ds, or even rhinoviruses could thus be found in the sequencing data generated from EV-Cs cell cultute supernatants co-infected by members of these species. Nonetheless, as our bioinformatics analyses was stringent enough to discriminate reads from two EV-Cs belonging to the same type, they are likely to discriminate reads from more divergent viruses, thereby

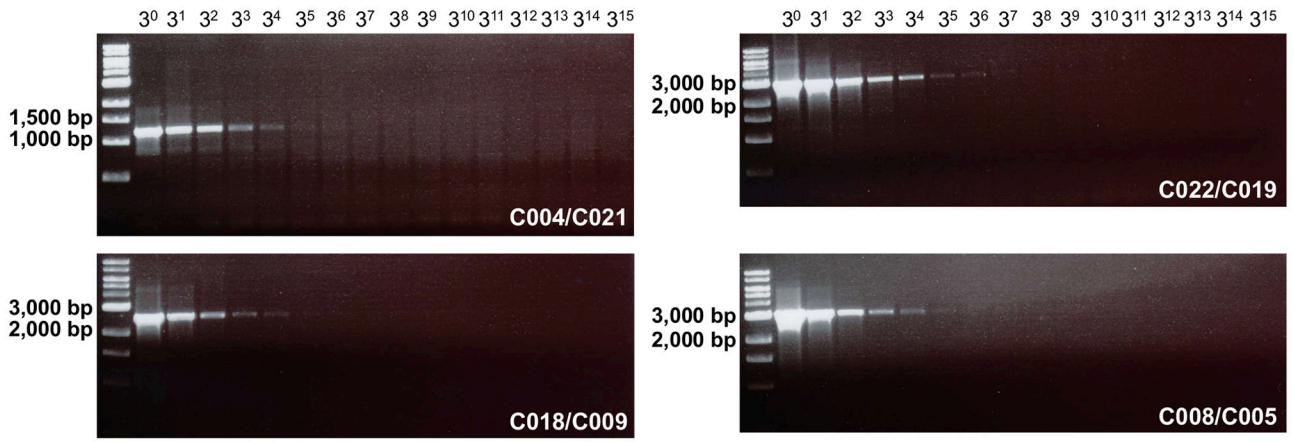


FIGURE 3 | Sensitivity of the different primer pairs on CV-A13 67001 RNA extracted from three-fold serial dilutions. The DNA size scale, expressed in base pairs (bp), is indicated on the left side of the gels.

TABLE 5 | Results obtained after sequencing and *de novo* assembly of mixtures of two viruses, CV-A13 67900 and CV-A13 67001.

Viral titer (log_{10} $TCID_{50}$)			Total number of reads	*De novo* assembly			
67900	67001	Ratio		67900		67001	
				Reads (%[a])	Average depth	Reads (%[b])	Average depth
7.5	7.5	1:1	267,226	84,674 (31%)	1540	178,860 (66%)	3396
7.5	7.2	2:1	224,848	94,097 (41%)	1689	121,239 (53%)	2313
7.5	6.9	4:1	236,342	129,959 (54%)	2105	104,977 (44%)	1999
7.5	6.5	10:1	147,338	98,121 (66%)	1892	38,977 (26%)	740

[a] *These percentages indicate the proportion of total reads that were included in the CV-A13 67900 contig through de novo assembly.*
[b] *These percentages indicate the proportion of total reads that were included in the CV-A13 67001 contig through de novo assembly.*

preventing the generation of contigs made of reads from viruses belonging to different species.

Our method was able to generate sequencing data even when no bands were observed on agarose gel after the RT-PCR step. This is a great advantage compared to Sanger-based sequencing that requires substantial amounts of DNA. Since the sensitivity of high-throughput sequencing depends as much on the sequencing depth (i.e., the total number of reads obtained for a given sample) as it does on the amount of virus genome copies in the sample, increasing the sequencing depth would overcome the low yield of RT-PCR amplification that could be observed for some samples. Another advantage of high-throughput sequencing compared to Sanger-based technics is that no gel-purification of the RT-PCR products was required, even in the presence of contaminating bands. In case of numerous samples to be analyzed, the RT-PCR products could be purified faster by using ultrafiltration-based or silica-based 96-well plates rather than individual silica-based columns.

For all samples, the coverage depth was heterogeneous along the genome. Heterogeneous coverage depths are often observed after high-throughput sequencing, partly because of the processing of the samples (Head et al., 2014; van Dijk et al., 2014). Thus, biases can be introduced by the random shearing of the DNA being sequenced (Poptsova et al., 2014) and by the PCR amplification performed during the library preparation (Aird et al., 2011). After these steps, some genomic regions can be overrepresented whereas others are underrepresented in the final libraries. Therefore, the coverage depth along the genome does not reflect necessarily the relative abundance of DNA amplicons in the original sample. However, in our experiments, the low coverage depth observed in some genomic regions did not impair the generation of the full-length consensus sequence of the corresponding genomes through *de novo* assembly.

In conclusion, we developed a set of generic primers for the synthesis of RT-PCR products that span the whole genome of EV-C isolates. This method was evaluated by using a panel of viruses already characterized by Sanger-based methods to allow the comparison of the sequences generated by our technique with those obtained previously. For *de novo* assembly, the raw data were analyzed in real conditions, i.e., with no reference to the sequences already obtained by Sanger-based methods. After assembly, the consensus sequences generated in this way were identical to the Sanger consensus sequences.

In this work, the RT-PCR products were sequenced by the Illumina sequencing technology but could be sequenced

on any high-throughput sequencing platform. Thus, the sequencing of DNA amplicons covering the whole genome of enterovirus A71 isolates on an Ion Torrent Personal Genomic Machine System was previously reported (Baronti et al., 2015).

This method will serve as a multipurpose tool for laboratories involved in the enterovirus surveillance. It could be used to quickly determine full-length genomic sequences of EV-Cs isolated in cell cultures during environmental surveillance or from patients. In particular, characterization of the genome of vaccine-derived PV, which generally display recombinant genomes made of PV and non-PV genetic sequences (Combelas et al., 2011), could be achieved quickly. The method could also be used to get the full-length sequences of EV-Cs belonging to uncharacterized collections, for example those constituted by the laboratories involved in diagnosis or environmental surveillance (Zaidi et al., 2016). Analyzing large panels of full-length EV-C genomes originating from such collections would help to describe the recombination events that occur between co-circulating viruses and to better understand how recombination drives EV-C evolution.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: MB and FD. Performed the experiments: MB, SS, MJ, PP. Analyzed the data: MB. Contributed reagents/materials/analysis tools: RR, RV, NJ. Wrote the manuscript: MB, FD. Critical revision: MA, BB, NJ.

FUNDING

This work was supported by the *Institut Pasteur* (PTR 484), the *Fondation Total,* and the US Department of Health and Human Services (grant No. 5 IDSEP140020-02-00). Patsy Polston is granted by a Pasteur Foundation's Gillings Pasteur Fellowship.

ACKNOWLEDGMENTS

The authors are indebted to Laura Brinas, Andreea Alexandru Maud Vanpeene, Sobhy Wilhame, and Vincent Enouf (Institut Pasteur, Pasteur International Bioresources network (PIBnet), Plateforme de microbiologie mutualisée (P2M), Paris, France) for performing the sequencing experiments.

REFERENCES

Aird, D., Ross, M. G., Chen, W. S., Danielsson, M., Fennell, T., Russ, C., et al. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* 12:R18. doi: 10.1186/gb-2011-12-2-r18

Anonymous (2004). "Titration of laboratory quality control standard," in *Polio Laboratory Manual,* ed WHO (Geneva: World Health Organization), 75–78.

Baronti, C., Piorkowski, G., Leparc-Goffart, I., de Lamballerie, X., and Dubot-Peres, A. (2015). Rapid next-generation sequencing of dengue, EV-A71 and RSV-A viruses. *J. Virol. Methods* 226, 7–14. doi: 10.1016/j.jviromet.2015.09.004

Berthet, N., Reinhardt, A. K., Leclercq, I., van Ooyen, S., Batejat, C., Dickinson, P., et al. (2008). Phi29 polymerase based random amplification of viral RNA as an alternative to random RT-PCR. *BMC Mol. Biol.* 9:77. doi: 10.1186/1471-2199-9-77

Bessaud, M., Jegouic, S., Joffret, M. L., Barge, C., Balanant, J., Gouandjika-Vasilache, I., et al. (2008). Characterization of the genome of human enteroviruses: design of generic primers for amplification and sequencing of different regions of the viral genome. *J. Virol. Methods* 149, 277–284. doi: 10.1016/j.jviromet.2008.01.027

Bessaud, M., Joffret, M. L., Holmblat, B., Razafindratsimandresy, R., and Delpeyroux, F. (2011). Genetic relationship between cocirculating Human enteroviruses species C. *PLoS ONE* 6:e24823. doi: 10.1371/journal.pone.0024823

Blomqvist, S., and Roivainen, M. (2016). Isolation and characterization of enteroviruses from clinical samples. *Methods Mol. Biol.* 1387, 19–28. doi: 10.1007/978-1-4939-3292-4_3

Brown, B., Oberste, M. S., Maher, K., and Pallansch, M. A. (2003). Complete genomic sequencing shows that polioviruses and members of human enterovirus species C are closely related in the noncapsid coding region. *J. Virol.* 77, 8973–8984. doi: 10.1128/JVI.77.16.8973-8984.2003

Caro, V., Guillot, S., Delpeyroux, F., and Crainic, R. (2001). Molecular strategy for 'serotyping' of human enteroviruses. *J. Gen. Virol.* 82(Pt 1), 79–91. doi: 10.1099/0022-1317-82-1-79

Combelas, N., Holmblat, B., Joffret, M. L., Colbere-Garapin, F., and Delpeyroux, F. (2011). Recombination between poliovirus and coxsackie A viruses of species C: a model of viral genetic plasticity and emergence. *Viruses* 3, 1460–1484. doi: 10.3390/v3081460

Cordey, S., Gerlach, D., Junier, T., Zdobnov, E. M., Kaiser, L., and Tapparel, C. (2008). The cis-acting replication elements define human enterovirus and rhinovirus species. *RNA* 14, 1568–1578. doi: 10.1261/rna.1031408

Cruz, C. D., Torre, A., Troncos, G., Lambrechts, L., and Leguia, M. (2016). Targeted full-genome amplification and sequencing of dengue virus types 1-4 from South America. *J. Virol. Methods* 235, 158–167. doi: 10.1016/j.jviromet.2016.06.001

Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., et al. (2008). Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. doi: 10.1186/1471-2164-9-5

Ge, F., Parker, J., Chul Choi, S., Layer, M., Ross, K., Jilly, B., et al. (2015). Preferential amplification of pathogenic sequences. *Sci. Rep.* 5:11047. doi: 10.1038/srep11047

Hall, R. J., Wang, J., Todd, A. K., Bissielo, A. B., Yen, S., Strydom, H., et al. (2014). Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J. Virol. Methods* 195, 194–204. doi: 10.1016/j.jviromet.2013.08.035

Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R., et al. (2014). Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* 56, 61–77. doi: 10.2144/000114133

Junttila, N., Lévêque, N., Magnius, L. O., Kabue, J. P., Muyembe-Tamfum, J. J., Maslin, J., et al. (2015). Complete coding regions of the prototypes enterovirus B93 and C95: phylogenetic analyses of the P1 and P3 regions of EV-B and EV-C strains. *J. Med. Virol.* 87, 485–497. doi: 10.1002/jmv.24062

Knowles, N., Hovi, T., Hyypiä, T., King, A. M. Q., Lindberg, A. M., Pallansch, M., et al. (2012). "Picornaviridae," in *Ninth Report of the International Committee on Taxonomy of Viruses*, eds A. M. Q. King, M. J. Adams, E. B. Carstens, and E. J. Lefkowitz (Amsterdam: Elsevier Academic Press), 855–880.

Kyriakopoulou, Z., Pliaka, V., Amoutzias, G. D., and Markoulatos, P. (2015). Recombination among human non-polio enteroviruses: implications for epidemiology and evolution. *Virus Genes* 50, 177–188. doi: 10.1007/s11262-014-1152-y

Monpoeho, S., Dehée, A., Mignotte, B., Schwartzbrod, L., Marechal, V., Nicolas, J. C., et al. (2000). Quantification of enterovirus RNA in sludge samples using single tube real-time RT-PCR. *Biotechniques* 29, 88–93. Available online at: http://www.biotechniques.com/BiotechniquesJournal/2000/July/Quantification-of-Enterovirus-RNA-in-Sludge-Samples-Using-Single-Tube-Real-Time-RT-PCR/biotechniques-44174.html

Mount, D. W. (2007). Using the Basic Local Alignment Search Tool (BLAST). *CSH Protoc.* 2007:pdb top17. doi: 10.1101/pdb.top17

Nelson, C. W., and Hughes, A. L. (2015). Within-host nucleotide diversity of virus populations: insights from next-generation sequencing. *Infect. Genet. Evol.* 30, 1–7. doi: 10.1016/j.meegid.2014.11.026

Nix, W. A., Oberste, M. S., and Pallansch, M. A. (2006). Sensitive, seminested PCR amplification of VP1 sequences for direct identification of all enterovirus serotypes from original clinical specimens. *J. Clin. Microbiol.* 44, 2698–2704. doi: 10.1128/JCM.00542-06

Poptsova, M. S., Il'icheva, I. A., Nechipurenko, D. Y., Panchenko, L. A., Khodikov, M. V., Oparina, N. Y., et al. (2014). Non-random DNA fragmentation in next-generation sequencing. *Sci. Rep.* 4:4532. doi: 10.1038/srep04532

Quiñones-Mateu, M. E., Avila, S., Reyes-Teran, G., and Martinez, M. A. (2014). Deep sequencing: becoming a critical tool in clinical virology. *J. Clin. Virol.* 61, 9–19. doi: 10.1016/j.jcv.2014.06.013

Rakoto-Andrianarivelo, M., Guillot, S., Iber, J., Balanant, J., Blondel, B., Riquet, F., et al. (2007). Co-circulation and evolution of polioviruses and species C enteroviruses in a district of Madagascar. *PLoS Pathog.* 3:e191. doi: 10.1371/journal.ppat.0030191

Rakoto-Andrianarivelo, M., Rousset, D., Razafindratsimandresy, R., Chevaliez, S., Guillot, S., Balanant, J., et al. (2005). High frequency of human enterovirus species C circulation in Madagascar. *J. Clin. Microbiol.* 43, 242–249. doi: 10.1128/JCM.43.1.242-249.2005

Rosseel, T., Van Borm, S., Vandenbussche, F., Hoffmann, B., van den Berg, T., Beer, M., et al. (2013). The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. *PLoS ONE* 8:e76144. doi: 10.1371/journal.pone.0076144

Sadeuh-Mba, S. A., Bessaud, M., Massenet, D., Joffret, M. L., Endegue, M. C., Njouom, R., et al. (2013). High frequency and diversity of species C enteroviruses in cameroon and neighboring countries. *J. Clin. Microbiol.* 51, 759–770. doi: 10.1128/JCM.02119-12

Savolainen-Kopra, C., and Blomqvist, S. (2010). Mechanisms of genetic variation in polioviruses. *Rev. Med. Virol.* 20, 358–371. doi: 10.1002/rmv.663

Tapparel, C., Siegrist, F., Petty, T. J., and Kaiser, L. (2013). Picornavirus and enterovirus diversity with associated human diseases. *Infect. Genet. Evol.* 14, 282–293. doi: 10.1016/j.meegid.2012.10.016

Thomson, E., Ip, C. L., Badhan, A., Christiansen, M. T., Adamson, W., Ansari, M. A., et al. (2016). Comparison of next generation sequencing technologies for the comprehensive assessment of full-length hepatitis C viral genomes. *J. Clin. Microbiol.* doi: 10.1128/JCM.00330-16. [Epub ahead of print].

Tokarz, R., Haq, S., Sameroff, S., Howie, S. R., and Lipkin, W. I. (2013). Genomic analysis of coxsackieviruses A1, A19, A22, enteroviruses 113 and 104: viruses representing two clades with distinct tropism within enterovirus C. *J. Gen. Virol.* 94(Pt 9), 1995–2004. doi: 10.1099/vir.0.053462-0

van Dijk, E. L., Jaszczyszyn, Y., and Thermes, C. (2014). Library preparation methods for next-generation sequencing: tone down the bias. *Exp. Cell Res.* 322, 12–20. doi: 10.1016/j.yexcr.2014.01.008

Wright, C. F., Morelli, M. J., Thebaud, G., Knowles, N. J., Herzyk, P., Paton, D. J., et al. (2011). Beyond the consensus: dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing. *J. Virol.* 85, 2266–2275. doi: 10.1128/JVI.01396-10

Zaidi, S. S., Asghar, H., Sharif, S., and Alam, M. M. (2016). Poliovirus laboratory based surveillance: an overview. *Methods Mol. Biol.* 1387, 11–18. doi: 10.1007/978-1-4939-3292-4_2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Bessaud, Sadeuh-Mba, Joffret, Razafindratsimandresy, Polston, Volle, Rakoto-Andrianarivelo, Blondel, Njouom and Delpeyroux. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 April 2016
doi: 10.3389/fmicb.2016.00543

Comparative Genetic Analyses of Human Rhinovirus C (HRV-C) Complete Genome from Malaysia

*Yam Sim Khaw[1], Yoke Fun Chan[2], Faizatul Lela Jafar[2], Norlijah Othman[3] and Hui Yee Chee[1]**

[1] Department of Medical Microbiology and Parasitology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Malaysia, [2] Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, [3] Department of Paediatrics, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Malaysia

Human rhinovirus-C (HRV-C) has been implicated in more severe illnesses than HRV-A and HRV-B, however, the limited number of HRV-C complete genomes (complete 5′ and 3′ non-coding region and open reading frame sequences) has hindered the in-depth genetic study of this virus. This study aimed to sequence seven complete HRV-C genomes from Malaysia and compare their genetic characteristics with the 18 published HRV-Cs. Seven Malaysian HRV-C complete genomes were obtained with newly redesigned primers. The seven genomes were classified as HRV-C6, C12, C22, C23, C26, C42, and pat16 based on the VP4/VP2 and VP1 pairwise distance threshold classification. Five of the seven Malaysian isolates, namely, 3430-MY-10/C22, 8713-MY-10/C23, 8097-MY-11/C26, 1570-MY-10/C42, and 7383-MY-10/pat16 are the first newly sequenced complete HRV-C genomes. All seven Malaysian isolates genomes displayed nucleotide similarity of 63–81% among themselves and 63–96% with other HRV-Cs. Malaysian HRV-Cs had similar putative immunogenic sites, putative receptor utilization and potential antiviral sites as other HRV-Cs. The genomic features of Malaysian isolates were similar to those of other HRV-Cs. Negative selections were frequently detected in HRV-Cs complete coding sequences indicating that these sequences were under functional constraint. The present study showed that HRV-Cs from Malaysia have diverse genetic sequences but share conserved genomic features with other HRV-Cs. This genetic information could provide further aid in the understanding of HRV-C infection.

Keywords: Malaysia, complete genome, HRV-C, genomic feature, phylogenetic, negative selection

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University Graduate School, Japan

Reviewed by:
Hirotaka Ode,
National Hospital Organization Nagoya Medical Center, Japan
Frederick Joseph Fuller,
North Carolina State University, USA
Reena Ghildyal,
University of Canberra, Australia

****Correspondence:***
Hui Yee Chee
cheehy@upm.edu.my

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *29 January 2016*
Accepted: *04 April 2016*
Published: *29 April 2016*

Citation:
Khaw YS, Chan YF, Jafar FL, Othman N and Chee HY (2016) Comparative Genetic Analyses of Human Rhinovirus C (HRV-C) Complete Genome from Malaysia. Front. Microbiol. 7:543. doi: 10.3389/fmicb.2016.00543

INTRODUCTION

Human rhinovirus (HRV) has been recognized as one of the respiratory viruses associated with common cold in humans since the late 1950s. HRV infection is a significant burden that is frequently associated with morbidity and mortality among hospitalized pediatric patients not only in Malaysia, but also worldwide (Etemadi et al., 2013). HRV has more than 100 serotypes, which have been clustered into three groups: HRV-A, B, and C (McIntyre et al., 2013). To date, HRV-A consists of 80 serotypes, HRV-B consists of 32 types, and HRV-C consists of 55 types (Picornaviridae Study Group, 2014).

HRV-C, which was discovered in 2006, is a new species of HRV and has led to the reemergence of research interest in HRV (Lamson et al., 2006; Lau et al., 2007). The unique characteristics of

HRV-C that distinguish it from HRV-A and HRV-B include the lack of usage of both major and minor receptors for host binding (Bochkov et al., 2011), potential resistance to pleconaril, a higher GC content, and a series of nucleotide deletions in the VP1 region, which yields a shorter complete genome (Lau et al., 2007).

A clearer picture of biological properties of HRV-A and HRV-B has been revealed through several studies that utilized a large number of HRV-A ($n > 80$) and HRV-B ($n > 30$) complete genome sequences in the Genbank (Palmenberg et al., 2009, 2010). Currently, there are only 18 HRV-C complete genomes (with complete 5′ non-coding region (NCR), 3′ NCR and open reading frame (ORF) sequences) available. This has greatly hampered the further study of this virus. Only partial and complete HRV-C VP1 and VP4/VP2 sequences have been largely published in the Genbank, hence, no in-depth analysis of the genomic features is possible.

Many studies have reported higher occurrence of HRV-C than HRV-A and HRV-B in severe respiratory diseases (McErlean et al., 2007; Miller et al., 2009). In addition, the association of specific HRV sequences with clinical severity is not well understood. To date, most of the published studies focused on the HRV-C capsid regions. Thus, a comparative sequence analysis using HRV-C complete genomes is an effective approach to gain more insight into this virus. In the present study, seven HRV-C complete genomes from children with respiratory tract infections in Malaysia were sequenced and comparative analyses of Malaysian isolates with other HRV-Cs were performed. This is the first report on the complete genome sequencing and comparative analyses of HRV-C from Malaysia. Seven HRV-C complete genome sequences were successfully amplified in the present study and comparative genetic analyses demonstrated that all HRV-Cs, as well as Malaysian isolates, share the same genomic features even though their nucleotide sequences are highly variable. This genetic information is useful for future studies of HRV-C pathogenesis and undoubtedly will make a significant contribution in the understanding of rhinovirus genomics.

MATERIALS AND METHODS

Patient Recruitment and Sample Collection

Seven nasopharyngeal aspirate (NPA) samples were collected from children with respiratory tract infection at University Malaya Medical Center (UMMC) in Malaysia between October 2010 and May 2011 (Medical Ethic Approval no. 788.3) (Chan et al., 2012). These samples were previously confirmed as HRV-C based on the sequence of the VP4/VP2 region (Chan et al., 2012). RNA was extracted using QIAamp Viral RNA Mini Kit (Qiagen, Germany) according to the manufacturer's instructions.

Redesigning Primers

Three HRV-C complete genomes (accession number: EF582385, EF582386, and EF582387) were aligned by using ClustalW implemented in the MEGA6 program package (Tamura et al., 2013). Primers were modified and redesigned based on the primers used by Lau et al. (2007), and the redesigned primers were based on conserved regions to generate amplicon of less than 1 kb (Supplementary Table 1).

Complete Genome Sequencing

cDNA was synthesized using the SuperScript III kit (Invitrogen, USA) according to the manufacturer's instructions with minor modifications. A combined random priming and oligo (dT) priming strategy was utilized to produce cDNA. The reaction mixture contained 2.75 μL of RNA, a mixture of 5 μM random hexamer (Thermo Scientific, USA) and 0.5 μM oligo-dT_{18} (Thermo Scientific, USA) (a ratio of 10:1), 0.5 mM dNTP, 1X first strand buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 5 mM DTT, 10 U RNaseOUT (Invitrogen, USA), 50 U SuperScript III reverse transcriptase, and nuclease-free water to a final volume of 5 uL. The mixture was incubated at 25°C for 10 min, followed by incubation at 55°C for 60 min, and a final incubation at 70°C for 15 min.

PCR amplification was performed with 1 μL cDNA, 1X Colorless GoTaq Flexi Buffer, 3 mM $MgCl_2$, 0.2 mM dNTP (Fermentas, USA), 1.0 μM of each forward and reverse primer, and 1 U GoTaq Flexi DNA Polymerase (Promega, USA) for a final volume of 25 μL. This reaction was subjected to thermal cycling at 95°C for 5 min, followed by 35 cycles each consisting of denaturation at 95°C for 1 min, annealing for 1 min (annealing temperature varies with different primers in Supplementary Table 1), and extension at 72°C for 1 min and a final extension at 72°C for 5 min using MyCycler Thermal Cycler (BioRad, USA). The 5′ and 3′ NCRs of the viral genome were amplified using the 5′ rapid amplification of cDNA ends (RACE) system and 3′ RACE system (Invitrogen, USA), respectively, according to manufacturer's protocol.

The PCR products were gel-purified using the gel/PCR DNA fragments extraction kit (Geneaid, Taiwan). These products were bi-directionally sequenced with both forward and reverse primers by First Base Laboratories Sdn Bhd (Selangor, Malaysia). NCRs and the insufficient concentration of purified PCR products were cloned into the yT&A vector (Yeastern, Taiwan) and transformed into Top 10F' *Escherichia coli* following the manufacturer's protocol. Then, plasmids from three positive bacteria clones were extracted and purified using the High-Speed Plasmid Mini Kit (Geneaid, Taiwan). These positive plasmids were sent to First Base Laboratories Sdn Bhd (Selangor, Malaysia) for sequencing using ABI3770 sequencer (Applied Biosystems, USA).

Genome Alignment

The sequences were edited and assembled with the use of BioEdit version 7.1.11 software (Hall, 1999). Eighteen HRV-C complete genome references (with complete 5′ NCR, 3′ NCR and ORF) up until January 2016 were retrieved from the GenBank. These HRV-Cs were used in comparative analysis and are referred to as "other HRV-Cs." Alignment of the seven HRV-C gene sequences completed in this study and the other HRV-Cs complete genomes was performed by using ClustalW implemented in the MEGA6 program package (Tamura et al., 2013).

Phylogenetic Analysis

JmodelTest 2.1.1 (Darriba et al., 2012) was used to determine the best-fit nucleotide substitution model. The maximum likelihood phylogenetic tree was constructed using General Time Reversible (GTR) as the best substitution model with gamma distributed

and invariant sites. To construct the phylogenetic trees, bootstrap values calculated from 1000 trees were implemented in the MEGA6 program package (Tamura et al., 2013). The bootstrap values of 70 or higher were considered as significant clustering.

Pairwise Similarity

Pairwise similarity among the nucleotide of HRV-Cs complete genomes and complete coding sequences, and complete coding deduced amino acid sequences were calculated with the use of the GeneDoc 2.7.000 software (Nicholas et al., 1997).

Sequence Analyses

The nucleotide sequences of Malaysian isolates and other HRV-Cs were translated into deduced amino acid sequences using MEGA6 program package (Tamura et al., 2013). The deduced amino acid and NCR sequences of Malaysian isolates and other HRV-Cs were aligned with reference sequences: HRV-16 (accession number: L24917, HRV-A), HRV-2 (accession number: X02316, HRV-A), HRV-1 (accession number: FJ445111, HRV-A), HRV-1B (accession number: D00239, HRV-A), HRV-3 (accession number: EF173422, HRV-B), HRV-14 (accession number: NC_001490, HRV-B), HRV-70 (accession number: DQ473489, HRV-B), or coxsackievirus B3 (accession number: M33854, CV-B3) using ClustalW implemented in the MEGA6 program package (Tamura et al., 2013) to determine the functional motifs.

Selective Pressure Analysis

The synonymous (dS) and non-synonymous (dN) changes at every codon of all the HRV-C complete coding sequences, including seven Malaysian isolates ($n = 25$), were estimated using three different selective pressure analyses implemented in DataMonkey (http://www.datamonkey.org) (Pond and Frost, 2005). The analyses were single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and internal fixed effects likelihood (IFEL). These methods were conducted using the GTR model of nucleotide substitution and neighbor-joining method to determine the rate for dN and dS. In addition, the estimation of positive and negative selections was determined using *p*-value of 0.1. The synonymous rate exceeding the non-synonymous rate was considered as negative selection (dS > dN), while the positive selection was defined as when the non-synonymous rate exceeds the synonymous rate (dN > dS). Neutral selection was defined when the non-synonymous rate equals to the synonymous rate (dN = dS).

Nucleotide Sequence Accession Numbers

The complete genomes of seven Malaysian isolates were deposited in the GenBank under the accession numbers KF734978, KJ675505, KJ675506, KJ675507, KP890662, KP890663, and KP890664.

RESULTS

Malaysian Isolates Showed Highly Diverse Nucleotide Sequences

Complete genomes of seven HRV-Cs (1515-MY-10, 1570-MY-10, 3430-MY-10, 3805-MY-10, 7383-MY-10, 8097-MY-11, and 8713-MY-10) were successfully sequenced in this study. Similar to other HRV-Cs, the nucleotide length of the Malaysian isolates complete genomes ranged from 7087 to 7127 bp, with an ORF that encoded for a polyprotein of 2139–2155 amino acids. Complete 5′ NCR and 3′ NCR lengths were 611–640 bp and 41–52 bp, respectively. The GC content of all the isolates was 42.3–43.8% (Supplementary Table 2).

Malaysian isolates complete genomes and complete coding sequences showed nucleotide similarity of 63–81% among themselves and 63–96% with other HRV-Cs (**Table 1**). In addition, Malaysian isolates demonstrated complete coding deduced amino acids similarity of 66–91% among themselves and 65–99% with other HRV-Cs. These results show that immense differences in both nucleotide and deduced amino acids sequences were observed among Malaysian isolates. Low nucleotide pairwise sequence similarity with different types of the other HRV-Cs suggests that Malaysian HRV-Cs are genetically diverse.

Phylogenetic Analysis Based on All the Available Complete Genomes Showed a Different Clustering Pattern Compared to VP1 Gene

Phylogenetic analysis of HRV-C complete genomes showed no clear geographical or temporal segregation of the Malaysian isolates and other HRV-Cs (**Figure 1**). The analysis revealed that 1515-MY-10 was closely related to 026 from Hong Kong, which has been classified as HRV-C6, and the nucleotide similarity was as high as 96% for the complete genome (**Table 1**). In addition, isolate 7383-MY-10 was also clustered with a group of C6 from Hong Kong and China and one HRV-C3 (QPM) from Australia. Isolate 3430-MY-10 showed a similar grouping pattern with 7383-MY-10, which this isolate grouped with an additional closer strain, NY-074 (C7) from USA compared to 7383-MY-10, while 8097-MY-11 grouped with C51 lineage from China and USA. Besides, isolate 3805-MY-10, which was closely clustered with LZY101 (C12) from China, also showed 92% of nucleotide similarity in the complete genome (**Table 1**). Isolate 1570-MY-10 was grouped together with 3805-MY-10 and LZY101 (C12), whereas 8713-MY-10 demonstrated a closer relationship with 2536 (C41) from the USA (**Figure 1**). Despite the obvious clustering pattern, the nucleotide similarities between 7383-MY-10, 3430-MY-10, 8097-MY-11, 1570-MY-10, 8713-MY-10 and their respective HRV-C in the clustering were only 74–81% (**Table 1**).

To perform the phylogenetic analysis based on the VP1 gene, a phylogenetic tree was built using the HRV-C complete VP1 gene from the available sequences in the Genbank. While there were many VP1 sequences in the Genbank for a single HRV-C type, only one VP1 sequence was chosen to represent each HRV-C type. Similar to the tree drawn with complete genome, 1515-MY-10 and 3805-MY-10 were still grouped with C6 and C12, respectively (**Figures 1**, **2**); however, isolate 1570-MY-10 was later grouped with C42, 3430-MY-10 with C22, 7383-MY-10 with C1, 8097-MY-11 with C26, and 8713-MY-10 with C23 (**Figure 2**). The observed differences were due to the absence of C1, C22, C23, C26 and C42 complete genomes in Genbank.

TABLE 1 | Nucleotide and deduced amino acid similarity between Malaysian isolates and HRV-C references.

HRV-C	Type assignment	Similarity (%)																			
		1515-MY-10			1570-MY-10			8713-MY-10			3430-MY-10			7383-MY-10			3805-MY-10			8097-MY-11		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
1515-MY-10	C6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1570-MY-10	C42	64	64	66	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
8713-MY-10	C23	65	64	67	68	68	74	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
3430-MY-10	C22	75	75	85	64	64	66	64	64	67	–	–	–	–	–	–	–	–	–	–	–	–
7383-MY-10	pat16	81	81	91	64	64	66	64	64	67	75	75	85	–	–	–	–	–	–	–	–	–
3805-MY-10	C12	64	63	66	77	76	87	69	69	75	63	64	66	64	64	66	–	–	–	–	–	–
8097-MY-11	C26	63	64	66	64	64	68	66	66	70	64	65	67	63	64	66	64	64	68	–	–	–
KF958310/6331	C2	64	64	65	65	65	67	66	65	69	64	64	67	64	64	66	64	64	67	67	67	72
EF186077/QPM	C3	85	84	95	65	65	66	65	65	67	75	75	86	81	80	90	64	64	67	64	65	67
EF582385/024	C4	64	64	68	66	66	70	67	66	71	64	64	68	64	64	68	66	66	70	65	66	70
EF582386/025	C5	65	64	68	64	64	67	66	66	68	65	65	69	65	64	68	65	65	67	64	65	67
EF582387/026	C6	96	96	99	64	64	66	65	64	67	76	75	85	81	81	91	64	64	66	64	64	67
JF317016/LZ651	C6	94	94	98	64	64	66	64	64	67	75	75	85	81	80	91	64	64	66	64	64	66
DQ875932/NY-074	C7	77	77	87	63	63	65	64	64	67	77	75	85	77	77	87	64	64	66	63	64	66
GQ223227/N4	C8	64	65	67	71	71	78	68	68	74	64	64	67	64	64	67	71	71	78	64	65	68
GQ223228/N10	C9	64	64	66	64	65	67	65	65	69	63	63	66	63	64	66	64	64	68	71	71	78
EU840952/CL170085	C11	65	65	68	64	64	67	65	64	68	65	65	68	65	64	67	65	64	67	64	64	67
JF317017/LZY101	C12	63	64	66	77	77	87	69	69	75	64	64	66	64	64	66	92	92	98	64	65	68
JF317014/LZY79	C15	64	64	67	69	69	75	73	73	83	64	64	67	65	64	68	68	68	74	65	66	68
GU219984/W10	C15	64	64	68	69	69	75	73	73	83	64	64	68	65	65	68	68	68	75	65	66	68
JF317013/LZ269	C25	64	64	68	68	68	75	72	72	81	64	64	68	64	63	67	69	69	76	65	66	69
JN205461/WA823M02	C39	69	68	75	64	64	67	65	65	69	70	69	75	69	68	74	64	64	68	65	65	68
KF958311/2536	C41	65	65	69	68	68	74	74	74	83	65	65	68	65	64	68	68	69	75	66	66	69
JF317015/LZ508	C51	63	64	66	64	65	68	65	65	69	63	64	67	63	64	67	64	64	68	76	76	86
JX291115/JAL-1	C51	63	64	65	64	65	69	66	67	72	63	64	66	63	63	66	65	65	70	72	71	78

[a] *Complete genome nucleotide similarity.*

[b] *Complete coding nucleotide similarity.*

[c] *Complete coding deduced amino acid similarity.*

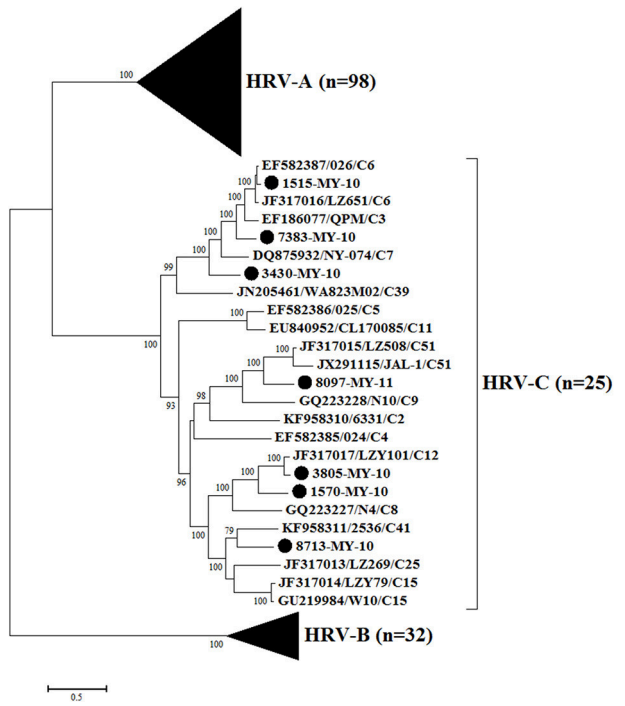


FIGURE 1 | Phylogenetic analysis of complete genome sequences of HRV-A, B, and C and seven Malaysian isolates used in this study. The Malaysian isolates were denoted as (●). Tree was constructed by using maximum likelihood method and General Time Reversible as substitution model with gamma distributed with invariant sites rate, with 1000 bootstrap replicates implemented in the program MEGA version 6. Black triangles represent HRV-A and HRV-B isolates, respectively. Only bootstrap values with higher than 70% were shown.

Pairwise Similarity Analysis Revealed Five New HRV-C Complete Genomes

Majority of the Malaysian isolates were consistent with pairwise distance threshold classification (Simmonds et al., 2010). They displayed more than 90% similarity with their respective HRV-C reference in both VP4/VP2 and VP1 nucleotide sequences, except for 7383-MY-10 (Supplementary Table 3). In a previous phylogenetic analysis, 7383-MY-10 demonstrated a closer relationship with pat16 in VP4/VP2 sequence (Chan et al., 2012). In this analysis, isolate 7383-MY-10 showed a highly similar result to pat16 with 93% in VP4/VP2 sequence, but did not show more than 90% similarity with any available HRV-C VP1 sequences at nucleotide level. The most similar HRV-C VP1 nucleotide sequence available in Genbank to 7383-MY-10 was HRV-C1 (NAT001, accession number: EF077279), for which the similarity was only 83%. Pat16 is one of the provisional HRV-Cs awaiting HRV-C type classification due to the absence of VP1 sequence (Picornaviridae Study Group, 2014).

Taken together the phylogeny and pairwise similarity analyses, the seven Malaysian isolates in this study, 1515-MY-10, 3805-MY-10, 3430-MY-10, 8713-MY-10, 8097-MY-11, 1570-MY-10, and 7383-MY-10 were classified as HRV-C6, C12,

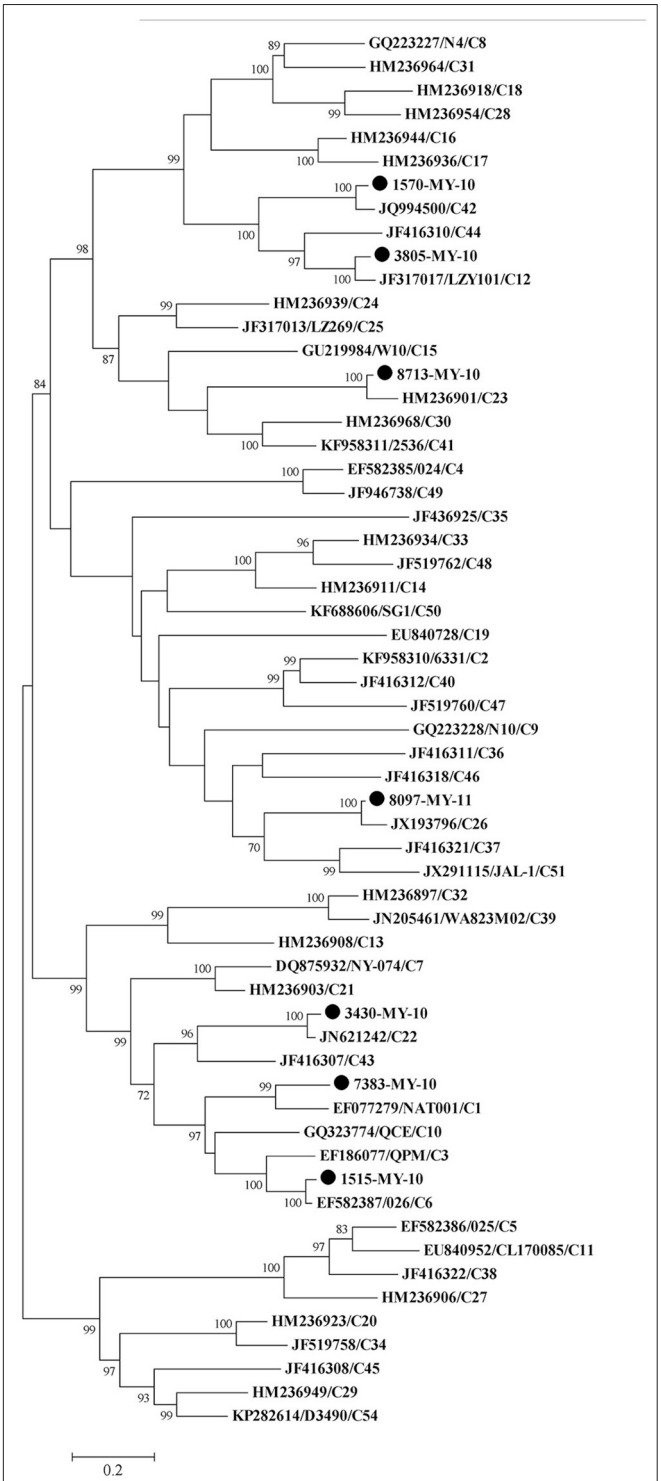


FIGURE 2 | Phylogenetic analysis of VP1 region of HRV-C and seven Malaysian isolates used in this study. The Malaysian isolates were denoted as (●). Tree was constructed using maximum likelihood method and General Time Reversible as substitution model with gamma, distributed with invariant sites rate with 1000 bootstrap replicates, implemented in the program MEGA version 6. Only bootstrap values of higher than 70% were shown.

C22, C23, C26, C42, and pat16, respectively. Since only two Malaysian isolates, 1515-MY-10 and 3805-MY-10 showed high complete genome sequence similarity with their respective HRV-C references, 026/C6 (accession number: EF582387) and LZY101/C12 (accession number: JF317017), therefore, the remaining five HRV-C isolates from Malaysia are likely to represent newly sequenced HRV-C complete genomes.

Malaysian Isolates Shared Similar Functional and Genetic Features with Other HRV-Cs

As HRV-C is not readily propagated in typical culture system, the immunogenic sites of HRV-C were predicted using comparative sequence analysis of HRV-A (HRV-2, accession number: X02316) (Appleyard et al., 1990) and HRV-B (HRV-14, accession number: K02121) (Rossmann et al., 1985). Similar to other HRV-Cs, the Malaysian isolates were shorter in length as compared to HRV-A and HRV-B due to several deletions in the VP1 loops (**Table 2**). These deleted regions are BC, DE, and HI loops. BC and DE loops are known as immunogenicity determinants in HRV-A and HRV-B, while HI loop is not related to immunogenic site; thus, HI loop is excluded from **Table 2**. The locations of immunogenic sites of HRV-A are different from HRV-B, except for position 85. Five (positions 85, 91, 95, 138, and 139) of six HRV-B immunogenic sites were deleted in HRV-Cs, whereas only two (positions 85 and 86) of nine HRV-A immunogenic sites were deleted in the Malaysian isolates and other HRV-Cs VP1 protein.

Malaysian isolates shared six out of seven and three out of nine conserved deduced amino acids with HRV-A and HRV-B major receptor binding sites, respectively (**Table 3**). The only dissimilar deduced amino acid of Malaysian isolates with HRV-A major receptor binding sites was located at position 180 in VP3 region (according to HRV-A). Malaysian isolates displayed several different deduced amino acids such as alanine (A), lysine (K), or serine (S) at position 180 in VP3 region, whereas HRV-A has proline (P) at that position. The other HRV-Cs shared five out of seven deduced amino acids with HRV-A major receptor binding sites (Lau et al., 2007). Intriguingly, an additional deduced amino acid in the HRV-A major receptor binding site, threonine (T), was displayed by isolate 8713-MY-10/C23 in the position 179 of VP3 region (according to HRV-A). On the other hand, Lys_{224} (K) amino acid in VP1 region was not seen in Malaysian isolates. The absence of this amino acid was due to the deletion of HI loop in the VP1 protein (data not shown).

Of the 25 amino acids, 15 and 14 amino acids corresponding to the pleconaril antiviral binding sites of HRV-A and HRV-B, respectively, were found in Malaysian isolates and other HRV-Cs (**Table 4;** Ledford et al., 2004). Also, these HRV-Cs demonstrated seven antiviral binding sites that were not analogous to HRV-A or HRV-B. These antiviral binding sites were located at positions 93, 113, 127, 166, 169, 221, and 243 (according to HRV-C024 accession number: EF582385). The putative antiviral binding sites of the Malaysian isolates and other HRV-Cs were superimposable onto the binding sites of HRV-A and HRV-B based on the same crucial C_α backbone present in each HRV group (Basta et al., 2014a). A minority group of HRV-Bs displaying Phe_{152} (F) and Leu_{191} (L) in the VP1 sequence were found to be associated with resistance to pleconaril (Ledford et al., 2004). Phe_{152} of HRV-B is equivalent to HRV-C Phe_{129}. Four Malaysian isolates (1515-MY-10/C6, 3430-MY-10/C22, 8713-MY-10/C23, and 7383-MY-10/pat16) and the majority of other HRV-Cs displayed only Phe_{129} amino acid but not Leu_{191} in the VP1 sequence (data not shown).

Many conserved motifs can be observed in the HRV genome. The genomic features of the Malaysian isolates and other HRV-Cs are summarized in **Table 5**. In the VP4 region, a conserved myristoylation motif, GAQVS motif, is essential for capsid protein assembly (Paul et al., 1987). This motif constitutes MGAQVS motif that is responsible for translation site of HRV polyprotein (Arden et al., 2010). Both of the

TABLE 2 | Summary of predicted HRV-C deduced amino acids and their positions in the VP1 region based on HRV-A (Appleyard et al., 1990) and -B immunogenic sites (Rossmann et al., 1985).

HRV	Position[#]													
	83	85[a]	86[a]	91[b]	92[a]	95[b]	138	139	264	267	268	269	276	278
HRV-A	–	T	L	–	E	–	–	–	K	D	R	S	T	P
HRV-B	Q	K	–	D	–	E	D	S	–	–	–	–	–	–
	Position*													
	83	NP[a]	NP[a]	NP[b]	84[a]	NP[b]	NP	NP	248	258	259	260	267	269
Other HRV-Cs	A/D/E/K/N/Q/R/S/T	Ø	Ø	Ø	D/E/N/Q/R/S	Ø	Ø	Ø	H/N	A/D/E/K/N/P/T/V	L/V	E/I/K/R/T	A/F/H/K/P/T/Y	A/D/E/G/N/P/Q/T
Malaysian isolates	D/K/N/P/R	Ø	Ø	Ø	E/N/Q/S/T	Ø	Ø	Ø	H/N	A/K/T/V	L	E/K/T	F/H/P/S/Y	D/E/N/R/T
Loop	–	BC	BC	BC	BC	BC	DE	DE	–	–	–	–	–	–

[#] *Position is according to reference sequence, HRV-A (HRV-2, accession number: X02316); HRV-B (HRV-14, accession number: K02121).*

[*] *Position is according to reference sequence, HRV-C024 (accession number: EF582385).*

[a] *Denotes Nim-1A for HRV-A &* [b] *Denotes Nim-1A for HRV-B, otherwise denotes Nim-1B for HRV-A and -B.*

(Ø), deletion.

TABLE 3 | Comparison of potential deduced amino acids in the VP1 and carboxy-terminal VP3 regions involved in HRV-A and B major receptor binding with Malaysian isolates and other HRV-Cs.

HRV	VP3					VP1			
	Amino acid position								
	179	**180**	**181**	**182**		**148**	**151**	**213**	
HRV-A major receptor	T[a]	P	D[b]	A/I/K/M/N[b]/S[b]/T		G[b]	A[b]/I/F/K/L/T[b]/V	A/D[b]/H	
	178	**179**	**180**	**181**		**133**	**136**	**200**	
Other HRV-Cs*	K/R	A/G/K/R/S	D[b]/S	D/E/K/N[b]/S[b]		G[b]	A[b]/G/H/P/S/T[b]/V	D[b]/G	
Malaysian isolates*	K/R/T[a]	A/K/S	D[b]	D/N[b]/S[b]		G[b]	A[b]/G/P/T[b]	D[b]	
	Amino acid position								
	177	**178**	**179**	**180**	**226**	**155**	**159**	**206**	**217**
HRV-B major receptor	D	K[b]/P	D[b]	T	Q/S	H/P[b]	K/N	H	V[c]
	178	**179**	**180**	**181**	**225**	**132**	**136**	**186**	**197**
Other HRV-Cs*	K/R	A/G/K[b]/R/S	D[b]/S	D/E/K/N/S	H/P	P[b]	A/G/H/P/S/T/V	A/I/K/L/N/R/S/T/V	A/E/L/P/Q/R/S/T/V[c]
Malaysian isolates*	K/R/T	A/K[b]/S	D[b]	D/N/S	P	P[b]	A/G/P/T	A/I/T/V	E/L/P/Q/S

Major receptor binding positions of HRV-A and HRV-B were described by Kolatkar et al. (1999).

**Position is according to reference sequence, HRV-C024 (accession number: EF582385).*

[a]Additional deduced amino acid similarity is shown by Malaysian isolate (8713-MY-10).

[b]Similar deduced amino acids shared by HRV-A and B major receptor with Malaysian isolates and other HRV-Cs.

[c]Additional deduced amino acid displayed by other HRV-Cs but not in Malaysian isolates.

motifs were identified in the Malaysian isolates and other HRV-Cs.

Malaysian isolates are likely to carry a chymotrypsin-like cysteine 2A protease with a catalytic triad, histidine (H), aspartic acid (D), and cysteine (C) overlapped with the GDCG motif that facilitates initial cleavage at its capsid polyprotein. It also causes shutdown of the host cell protein synthesis (Hughes and Stanway, 2000). A zinc ligand, C-C-C-H, that helps to bind a zinc ion firmly for enzymatic activities was identified in the 2A protein of the Malaysian isolates (Petersen et al., 1999). Two hydrophobic regions of the 2B protein were identified in the Malaysian isolates. The first hydrophobic region is postulated to form a cationic amphipathic alpha helix, which is the major determinant for permeabilization of a host's plasma membrane (Agirre et al., 2002). Meanwhile, the second hydrophobic region may serve as a transmembrane domain to inhibit host protein secretion (van Kuppeveld et al., 1997). In addition, a NTPase motif, crucial for NTP binding and composed of the GXPGXGKS sequence (Gorbalenya et al., 1988), was found in the 2C protein of the Malaysian isolates. The DDLXQ motif is crucial for putative helicase function of 2C protein (de Souza Luna et al., 2008). This motif was found in one of the Malaysian isolates (8713-MY-10/C23). The majority of the Malaysian isolates displayed the DDVXQ motif, while two of the Malaysian isolates (1570-MY-10/C42 and 3805-MY-10/C12) showed DDIXQ motif. It is postulated that these three motifs (DDLXQ, DDVXQ and DDIXQ) play the same role as the amino acid properties of L, V, and I are similar. In addition, the Malaysian isolates 2C protein showed a cysteine-rich motif (CX2-4CX6-8CX3-4C) resembling a zinc finger, which is important in viral RNA replication (Pfister et al., 2000). Overall, the features of the Malaysian isolates 2A, 2B, and 2C proteins are similar to those of the other HRV-Cs.

Since there are no three-dimensional structures of HRV-C 3A protein, the characteristics of HRV-C 3A protein were postulated based on another picornavirus, coxsackievirus. Hydrophobic packing and intermolecular salt bridges are important in coxsackievirus replication (Wessels et al., 2006). Malaysian isolates 3A protein displayed I-L-L-S-V instead of I-L-L-V-V, suggesting the absence of hydrophobic packing. In addition, no intermolecular salt bridge was identified in the Malaysian isolates 3A protein. Tyrosine, an important amino acid in the 3B protein for linking covalently with viral RNA through a phosphodiester bond (Ambros and Baltimore, 1978), was conserved in the Malaysian isolates. Similar to the 2A protein, 3C protein of Malaysian isolates was probably a chymotrypsin-like cysteine protease with slightly different catalytic triads, histidine (H), glutamic acid (E), and cysteine (C) overlapped with GQCG motif (Matthews et al., 1994). Additionally, a substrate binding pocket, GXH (Matthews et al., 1994), was found in the 3C protein of four of the Malaysian isolates (1515-MY-10/C6, 3430-MY-10/C22, 7383-MY-10/pat16, and 8097-MY-11/C26). Moreover, the Malaysian isolates 3C protein showed a conserved RNA binding motif, KFRDI (Shih et al., 2004). For the 3D protein, four conserved motifs; KDELR, GMPSG, YGDD, and FLKR (Appleby et al., 2005) were identified in Malaysian isolates. The KDELR motif constitutes part of a loop that joins the finger and thumb domain, which forms the active site of the 3D polymerase (Appleby et al., 2005). The YGDD motif is crucial for nucleotidyl transfer reaction (Beese and Steitz, 1991). Overall,

TABLE 4 | Comparison of majority deduced amino acids involved in antiviral binding against pleconaril for HRV-A and HRV-B with Malaysian isolates and other HRV-Cs.

HRV	VP1 region																								
	Position[#]																								
	98	**99**	**100**	**101**	**118**	**120**	**122**	**124**	**142**	**143**	**144**	**166**	**167**	**168**	**179**	**181**	**184**	**190**	**192**	**208**	**212**	**214**	**217**	**238**	**260**
HRV-A	I[a]	N[a]	L	Q[b]	F[a]	S	I[b]	L	Y	M[a]	Y[a]	A	S[a]	V[a]	F[a]	L	L	Y[a]	M[b]	T	N[a]	M[a]	L[b]	H	H
	Position*																								
	104	**105**	**106**	**107**	**124**	**126**	**128**	**130**	**150**	**151**	**152**	**174**	**175**	**176**	**186**	**188**	**191**	**197**	**199**	**215**	**219**	**221**	**224**	**245**	**267**
HRV-B	I[a]	N[a]	L	S	F[a]	S	Y	I[c]	A	M[a]	Y[a]	P[c]	S[a]	V[a]	F[a]	V	V	Y[a]	C	I[c]	N[a]	M[a]	M	H	G
	Position[@]																								
	91	**92**	**93**	**94**	**111**	**113**	**115**	**117**	**127**	**128**	**129**	**151**	**152**	**153**	**164**	**166**	**169**	**175**	**177**	**195**	**199**	**201**	**204**	**221**	**243**
Other HRV-Cs	I[a]	N[a]/S/T	F/Y	Q[b]	F[a]/Y	L/M	I[b]/V	I[c]/V	I/T/V	M[a]	F/Y[a]	P[c]	S[a]	I/V[a]	F[a]	I	M/T	Y[a]	I/M[b]/V	I[c]/T	D/N[a]	M[a]	L[b]	A/G/I/V	D/F/N/Y
Malaysian isolates	I[a]	N[a]/S	F	Q[b]	F[a]	L/M	I[b]/V	I[c]/V	I/T/V	M[a]	F/Y[a]	P[c]	S[a]	V[a]	F[a]	I	T	Y[a]	M[b]	I[c]	N[a]	M[a]	L[b]	A/I/V	F/Y

These antiviral binding sites were described by Ledford et al. (2004).

[a]*Similar deduced amino acids shared in HRV-A and HRV-B with Malaysian isolates and other HRV-Cs.*

[b]*Similar deduced amino acids shared in HRV-A with Malaysian isolates and other HRV-Cs.*

[c]*Similar deduced amino acids shared in HRV-B with Malaysian isolates and other HRV-Cs.*

[#]*Position is according to reference sequence, HRV-16 (accession number: L24917).*

**Position is according to reference sequence, HRV-14 (accession number: K02121).*

[@]*Position is according to reference sequence, HRV-C024 (accession number: EF582385).*

TABLE 5 | Summary of the genomic features of Malaysian isolates and other HRV-Cs.

Region	Conserved motif; Features	Characteristics	
		Malaysian isolates	Other HRV-Cs
VP4	Myristylation site: GAQVS	Yes	Yes
	Translation site: MGAQVS	Yes	Yes
2A	Chymotrypsin-like protease with catalytic triad: H-D-C and GXCG motif	Yes	Yes
	Zinc ligand: C-C-C-H	Yes	Yes except N10/C9
2B	Two hydrophobic regions	Yes	Yes
2C	NTPase motif: GXXGXGKS	Yes	Yes
	Helicase: DDLXQ	One of the seven Malaysian isolates	Seven of 18 other HRV-Cs
	Cysteine rich motif (CX2-4CX6-8CX3-4C)	Yes	Yes
3A	Hydrophobic packing I-L-L-V-V	No	No
	Intermolecular salt bridge D-K-K	No	No
3B	Phosphodiester bond: Y_3	Yes	Yes
3C	Chymotrypsin-like protease with catalytic triad: H-E-C and GQCG motif	Yes	Yes
	Substrate binding pocket: GXH	Four of the seven Malaysian isolates	Nine of 18 other HRV-Cs
	RNA binding motif: KFRDI	Yes	Yes
3D	KDELR motif	Yes	Yes
	GMPSG motif	Yes	Yes
	YGDD motif	Yes	Yes
	FLKR motif	Yes	Yes except JAL-1/C51
5′ NCR	Initial sequence, UU(A/G)AA(A/G)C(U/A)G	Yes	Yes
	Two pyrimidine rich segments	Yes	Yes
3′ NCR	Domain Y	Yes	Yes

the characteristics of the Malaysian isolates 3A, 3B, 3C, and 3D proteins are similar to those of the other HRV-Cs.

A nine bp relatively conserved nucleotides sequence, UU(A/G)AA(A/G)C(U/A)G, was identified in the initial sequence of each Malaysian isolate 5′ NCR, where a VPg protein binds to the first U of the sequence (Palmenberg et al., 2009) and initiates its translation. In addition, two pyrimidine-rich segments were observed in the Malaysian isolates. The first segment is postulated as a determinant of HRV-C pathogenecity as this segment is equivalent to neurovirulence tropism of poliovirus (Palmenberg et al., 2009); whereas, the second segment represents the HRV ribosome entry region which is important for translation initiation (Borman and Jackson, 1992). Malaysian isolates were found to share the same 5′ NCR characteristics with other HRV-Cs. Similar to other HRV-Cs, Malaysian isolates demonstrated domain Y in 3′ NCR (Pilipenko et al., 1992).

HRV-C Complete Coding Sequences were Predominated by Purifying Selective Pressure

A total of 2168 codon sites of the Malaysian isolates and other HRV-Cs aligned complete coding sequence were utilized for the selective pressure analysis. Negative selection was found to be the dominant selective pressure on HRV-C complete coding sequences using SLAC, FEL, and IFEL methods at $P < 0.1$ level. These analyses (SLAC, FEL, and IFEL) found that 81% (1758/2168), 87% (1886/2168), and 79% (1716/2168) codon sites, respectively, were highly significant for purifying selections. The IFEL method demonstrated three positive selected sites but SLAC and FEL methods did not demonstrate any. These three positive selected sites were located in 852, 943, and 1030 codon position (**Figure 3**).

DISCUSSION

Respiratory infection is a significant burden in Malaysia (Khor et al., 2012). HRV infection, although usually associated with mild symptoms, can be associated with more severe acute respiratory infections and complications including asthma exacerbations (Lau et al., 2007). Due to the higher incidence of HRV-C in causing severe diseases and insufficient availability of HRV-C complete genome sequences for further analyses, the present study aimed to sequence the complete genome of Malaysian HRV-Cs and to perform genetic analyses of these viruses. These samples were previously confirmed as HRV-C based on the phylogenetic analysis of the VP4/VP2 region and designated as 1515-MY-10, 1570-MY-10, 3430-MY-10, 3805-MY-10, 7383-MY-10, 8097-MY-11, and 8713-MY-10 (Chan et al., 2012). Seven HRV-C complete genome sequences consisting of 5′ and 3′ NCR, and the ORF were successfully amplified in this study. These seven Malaysian isolates were isolated from patients displaying a range of respiratory diseases such as bronchiolitis, staphylococcal pneumonia and acute exacerbations of bronchial asthma (AEBA). Low nucleotide similarity was observed in Malaysian isolates when compared to other HRV-Cs suggesting

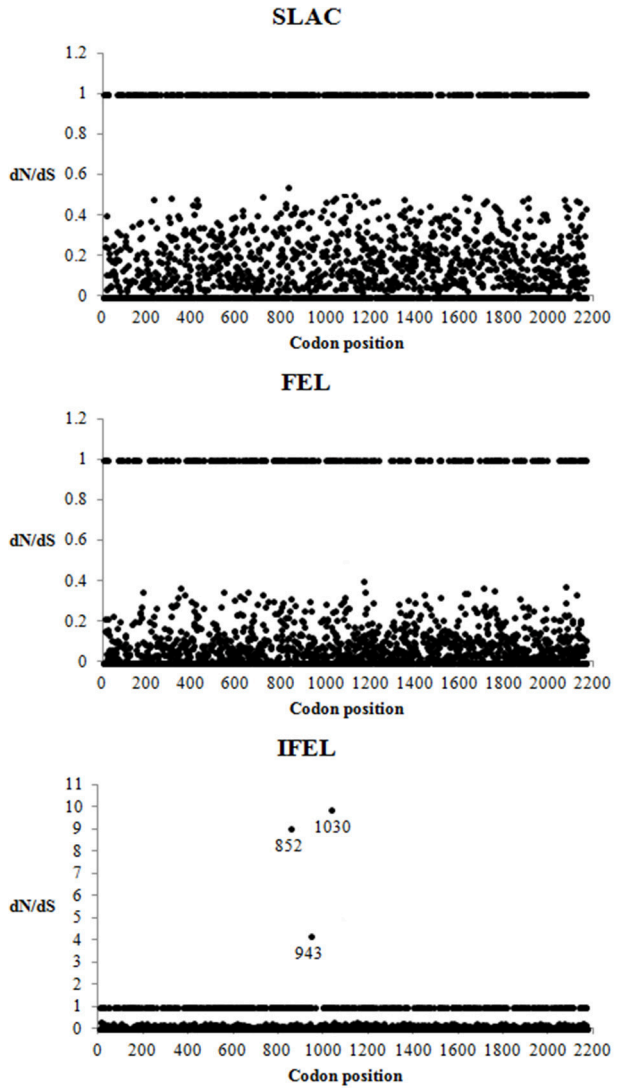


FIGURE 3 | Detection of selection pressures on complete coding sequences of HRV-C using SLAC, FEL, and REL methods. dN/dS > 1, dN/dS = 1, and dN/dS < 1 indicated codons under positive, neutral and negative selection, respectively. The codons with positive selection were labeled with their codon position.

that these isolates are genetically diverse. Similar genetic diversity has been reported in HRV-C as well as HRV-A and HRV-B (Palmenberg et al., 2009). In addition, the previous primer sets from Lau et al. (2007) were unable to amplify all of the Malaysian isolates, further demonstrating the diverse nucleotide sequences of Malaysian isolates.

Groupings in phylogenetic analysis of HRV-C complete genome were different from previous VP4/VP2 phylogenetic analysis (Chan et al., 2012) which clustered 1515-MY-10, 3805-MY-10, 3430-MY-10, 8713-MY-10, 8097-MY-11, 1570-MY-10, and 7383-MY-10 with HRV-C6, C12, C22, C23, C26, C42, and pat16, respectively. In the present study, phylogenetic analysis of VP1 gene demonstrated a similar clustering pattern with VP4/VP2 gene phylogenetic analysis, except for 7383-MY-10. 7383-MY-10 was clustered with C1 in the VP1 phylogenetic tree, and the similarity was only 83%. The similarity is inconsistent with the classification proposed by Simmonds et al. (2010) suggesting a paucity of representative VP1 sequences for pat16. On the other hand, several short VP1 and VP4/VP2 sequences of C22, C23, C26, and C42 are available in Genbank. These sequences showed at least 87% similarity with their respective Malaysian isolate sequences (1570-MY-10, 3430-MY-10, 8097-MY-11, and 8713-MY-10).

McErlean et al. (2007) determined that the length of HRV-C is shorter than HRV-A and HRV-B, as a result of several deletions in the BC, DE, and HI loops of VP1 region. Malaysian isolates have the same genome length as other HRV-Cs. The BC and DE loops are vital in major neutralization sites, especially Nim-1A and Nim-1B sites (Lau et al., 2007). The deletions of the putative neutralization sites in HRV-C remove the exposed loops and prevent the contact with host antibody; thus, HRV-C may have the possibility to escape host antibody recognition. This could be the reason why HRV-C is responsible for more severe respiratory diseases as compared to other HRVs (Basta et al., 2014b). Besides that, more deletions in HRV-C of HRV-B immunogenic sites were observed as compared with HRV-A sites, suggesting that a vaccine or antiviral targeting the neutralization sites for HRV-C could be designed based on HRV-A instead of HRV-B.

As for host receptor binding, Lau et al. (2007) reported that HRV-Cs shared five out of the seven and four out of the nine conserved residues with major receptors of HRV-A and HRV-B, respectively, based on a study by Kolatkar et al. (1999). The authors speculated that these HRV-Cs may not exploit the major receptor for host binding. One of the Malaysian isolates, 8713-MY-10/C23, showed that its putative major receptor binding sites were very similar to HRV-A major receptor binding sites, as compared to other HRV-Cs. We speculate that this Malaysian isolate may not utilize the HRV-A major receptor because the binding sites required three fully conserved amino acids (Thr_{179}, Pro_{180}, and Asp_{181}) in the VP3 region (Laine et al., 2006), and this was not conserved in this isolate. Lack of amino acid similarity between Malaysian isolates and HRV-B major receptor binding site also suggests that the Malaysian isolates may not exploit this receptor either. In addition, the deletion of the HI loop of VP1 gene in the Malaysian isolates has removed the Lys_{224} amino acid, which is crucial for minor receptor binding (Lau et al., 2007). Bochkov et al. (2011) demonstrated that HRV-C was still able to attach to the target cells although ICAM-1 (major receptor) and LDLR (minor receptor) antibodies were used to block these receptors. Taken together, these observations may imply that major and minor receptors are not utilized by HRV-C for attachment. Overall, the Malaysian isolates, like other HRV-Cs, are unlikely to utilize both the major and minor receptors for host binding based on genome sequence analysis. Recently, cadherin-related family member 3, a transmembrane protein with unknown function, was discovered as a possible receptor for HRV-C (Bochkov et al., 2015).

Arden et al. (2010) demonstrated 11 and 12 residue differences of HRV-C (HRVC-QCE) compared with HRV-A and HRV-B, respectively, among the 25 important contact residues of the

antiviral binding pocket. With the increasing number of HRV-Cs, the Malaysian isolates and other HRV-Cs demonstrated a higher number of similar residues with HRV-A and HRV-B with respect to the antiviral binding pocket. Basta et al. (2014a) showed that four unique antiviral residues, Phe/Tyr$_{96}$, Leu/Met$_{116}$, Ile/Val$_{130}$, and Ile$_{169}$ were only found in HRV-C, but not HRV-A and HRV-B. These residues were seen in the Malaysian isolates and are equivalent to Phe/Tyr$_{93}$, Leu/Met$_{113}$, Ile/Thr/Val$_{127}$, and Ile$_{166}$, respectively, in this study. Phe$_{96}$ and Met$_{116}$ of C15 have been shown to cause occasional steric clashes with some antiviral drugs (Basta et al., 2014a). According to the antiviral susceptibility pattern constructed by Basta et al. (2014a), Malaysian isolates and other HRV-Cs showed an additional unique residue, Ala/Gly/Ile/Val$_{221}$ in their sequences compared with HRV-A and HRV-B. These different residues of HRV-C undoubtedly changed the characteristic of HRV-C VP1 protein structure for antiviral binding, as compared to HRV-A and HRV-B. In an antiviral study, Ledford et al. (2004) demonstrated that apparent resistance to pleconaril was developed from HRV-B with Phe$_{152}$ and Leu$_{191}$ amino acids in the VP1 region. In the following year, Ledford et al. (2005) concluded that the resistance effect was most profound if both Phe$_{152}$ and Leu$_{191}$ were present concurrently in HRV. In the present study, none of the Malaysian isolates and other HRV-Cs demonstrated the simultaneous presence of these two amino acids. Recently, Hao et al. (2012) showed that HRV-C15 with only Phe$_{129}$ amino acid (same as HRV-B Phe$_{152}$ amino acid) alone was capable of confering noticeable resistance to pleconaril. Most of the Malaysian isolates and other HRV-Cs demonstrated the presence of Phe$_{129}$ amino acid in the VP1 sequence, suggesting that they are likely resistant to pleconaril.

Negative selection was the dominant selective pressure acting on the HRV-C complete coding sequences. This result is in agreement with Kistler et al. (2007), which reported that most of the HRV coding regions were under strong negative selective pressure. Besides that, most of the motifs in the complete coding sequences displayed low ratio of the number of non-synonymous substitutions per non-synonymous site to the number of synonymous substitutions per synonymous site. This indicates that these motifs have been under functional constraint throughout evolution. No positive selected sites were identified in HRV-C capsid region (Kuroda et al., 2015), whereas a positive selected site was reported in VP1 region of all HRVs complete coding sequences (86 HRV-A, 33 HRV-B, and 14 HRV-C) (Waman et al., 2014). A larger HRV-C dataset ($n = 25$) was exploited in the present study and several positive selective sites were identified using the IFEL method. One of these positive selected sites (852 codon position) was located in C-terminus of VP1 region, and it correlates with Waman et al. (2014). Another two positive selective pressures were found in 2A and 2B gene, respectively. The biological relevance of these three positive selected sites is hypothesized to increase in the HRV-C survival rate. For example, the positive selective site in C-terminus VP1 region is located on the outer surface, corresponding to the HRV-A antigenic site (269 position in HRV-C VP1 gene), and possibly interacts with host immune system (**Table 2**) resulting in the alteration of HRV-C immunogenicity. However, this hypothesis will require further support and verification from experimental data. For instance, the positively selected site is mapped within the Nim sites of HRV-A, it remains obscure whether this positive selective pressure affects the immunogenicity of HRV-C as the immunogenic determinants of HRV-C have not been identified. To our knowledge, no studies focusing on these selected sites in HRV-C have been reported. Future studies are necessary to determine the biological importance of these positive selected sites. In turn, these findings can identify the pathogenic role of HRV-C in causing severe diseases.

CONCLUSION

Complete genomes of seven HRV-C Malaysian isolates have been sequenced and were classified as C6, C12, C22, C23, C26, C42, and pat16 based on the pairwise distance threshold classification. This is the first report of complete genomes of C22, C23, C26, C42, and pat16. This study found that the putative biological characteristics of the Malaysian isolates were similar to those of other HRV-Cs. Moreover, negative selective pressure was the predominant pressure acting on complete coding sequences of HRV-C, and this pressure conserved the functional motif of these HRV-Cs, although the nucleotides were diverse. These indicate that a similar treatment and control of HRV-C can be applied to patients infected with these wide ranges of HRV-Cs. This study has represented the most up-to-date information about HRV-C genomics and shown that although HRV-Cs have diverse genetic sequences, they share conserved genomic features. Further *in vitro* and *in vivo* study will be required to clarify the HRV-C antiviral resistance and receptor binding sites that were predicted based on the genome sequences.

AUTHOR CONTRIBUTIONS

HYC, YFC, and NO participated in design of the study. YSK and FLJ performed the experiments. YSK, HYC, and YFC analyzed the data and drafted the manuscript. All authors revised and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Cassi Henderson for checking the English. This study was supported by Research University Grant Scheme (RUGS grant 9370300), Universiti Putra Malaysia and Exploratory Research Grant Scheme (ERGS grant 5527164) from Ministry of Education, Malaysia awarded to HYC and High Impact Research Grant E000013-20001, University Malaya awarded to YFC. The funders had no role in study design, data collection and analysis, decision to publish or preparation of manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.00543

REFERENCES

Agirre, A., Barco, A., Carrasco, L., and Nieva, J. L. (2002). Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *J. Biol. Chem.* 277, 40434–40441. doi: 10.1074/jbc.M205393200

Ambros, V., and Baltimore, D. (1978). Protein is linked to the 5′end of poliovirus RNA by a phosphodiester linkage to tyrosine. *J. Biol. Chem.* 253, 5263–5266.

Appleby, T. C., Luecke, H., Shim, J. H., Wu, J. Z., Cheney, W. I., Zhong, W. D., et al. (2005). Crystal structure of complete rhinovirus RNA polymerase suggests front loading of protein primer. *J. Virol.* 79, 277–288. doi: 10.1128/JVI.79.1.277-288.2005

Appleyard, G., Russell, S. M., Clarke, B. E., Speller, S. A., Trowbridge, M., and Vadolas, J. (1990). Neutralization epitopes of human rhinovirus type 2. *J. Gen. Virol.* 71, 1275–1282. doi: 10.1099/0022-1317-71-6-1275

Arden, K. E., Faux, C. E., O'Neill, N. T., McErlean, P., Nitsche, A., Lambert, S. B., et al. (2010). Molecular characterization and distinguishing features of a novel human rhinovirus (HRV) C, HRVC-QCE, detected in children with fever, cough and wheeze during 2003. *J. Clin. Virol.* 47, 219–233. doi: 10.1016/j.jcv.2010.01.001

Basta, H. A., Ashraf, S., Sgro, J. Y., Bochkov, Y. A., Gern, J. E., and Palmenberg, A. C. (2014a). Modeling of the human rhinovirus C capsid suggests possible causes for antiviral drug resistance. *Virology* 448, 82–90. doi: 10.1016/j.virol.2013.10.004

Basta, H. A., Sgro, J. Y., and Palmenberg, A. C. (2014b). Modeling of the human rhinovirus C capsid suggests a novel topography with insights on receptor preference and immunogenicity. *Virology* 448, 176–184. doi: 10.1016/j.virol.2013.10.006

Beese, L. S., and Steitz, T. A. (1991). Structural basis for the 3′-5′exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *EMBO J.* 10, 25–33.

Bochkov, Y. A., Palmenberg, A. C., Lee, W. M., Rathe, J. A., Amineva, S. P., Sun, X., et al. (2011). Molecular modeling, organ culture and reverse genetics for a newly identified human rhinovirus C. *Nat. Med.* 17, 627–633. doi: 10.1038/nm.2358

Bochkov, Y. A., Watters, K., Ashraf, S., Griggs, T. F., Devries, M. K., Jackson, D. J., et al. (2015). Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5485–5490. doi: 10.1073/pnas.1421178112

Borman, A., and Jackson, R. J. (1992). Initiation of translation of human rhinovirus RNA: mapping the internal ribosome entry site. *Virology* 188, 685–696. doi: 10.1016/0042-6822(92)90523-R

Chan, Y. F., Jafar, F. L., Nathan, A. M., de Bruyne, J. A., Hassan, A., Nor'e, S. S., et al. (2012). Diverse human rhinoviruses A and C from children with respiratory infections in Kuala Lumpur, Malaysia. *J. Infection* 64, 633–636. doi: 10.1016/j.jinf.2012.03.011

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9:772. doi: 10.1038/nmeth.2109

de Souza Luna, L. K., Baumgarte, S., Grywna, K., Panning, M., Drexler, J. F., and Drosten, C. (2008). Identification of a contemporary human parechovirus type 1 by VIDISCA and characterization of its full genome. *Virol. J.* 5:26. doi: 10.1186/1743-422X-5-26

Etemadi, M. R., Othman, N., Savolainen-Kopra, C., Sekawi, Z., Wahab, N., and Sann, L. M. (2013). Biodiversity and clinico-demographic characteristics of human rhinoviruses from hospitalized children with acute lower respiratory tract infections in Malaysia. *J. Clin. Virol.* 58, 671–677. doi: 10.1016/j.jcv.2013.05.017

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). A conserved NTP-motif in putative helicases. *Nature* 333:22. doi: 10.1038/333022a0

Hall, T. A. (1999). BioEdit, a user-friendly biological sequence alignment editor and analysis program Windows 95/98/NT. *Nucleic Acid Symp. Ser.* 41, 95–98. doi: 10.1021/bk-1999-0734.ch008

Hao, W., Bernard, K., Patel, N., Ulbrandt, N., Feng, H., Svabek, C., et al. (2012). Infection and propagation of human rhinovirus C in human airway epithelial cells. *J. Virol.* 86, 13524–13532. doi: 10.1128/JVI.02094-12

Hughes, P. J., and Stanway, G. (2000). The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation. *J. Gen. Virol.* 81, 201–207. doi: 10.1099/0022-1317-81-1-201

Khor, C. S., Sam, I. C., Hooi, P. S., Quek, K. F., and Chan, Y. F. (2012). Epidemiology and seasonality of respiratory viral infections in hospitalized children in Kuala Lumpur, Malaysia: a retrospective study of 27 years. *BMC Pediatr.* 12:32. doi: 10.1186/1471-2431-12-32

Kistler, A. L., Webster, D. R., Rouskin, S., Magrini, V., Credle, J. J., Schnurr, D. P., et al. (2007). Genome-wide diversity and selective pressure in the human rhinovirus. *Virol. J.* 4:40. doi: 10.1186/1743-422X-4-40

Kolatkar, P. R., Bella, J., Olson, N. H., Bator, C. M., Baker, T. S., and Rossmann, M. G. (1999). Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. *EMBO J.* 18, 6249–6259. doi: 10.1093/emboj/18.22.6249

Kuroda, M., Niwa, S., Sekizuka, T., Tsukagoshi, H., Yokoyama, M., Ryo, A., et al. (2015). Molecular evolution of the VP1, VP2, VP3 genes in human rhinovirus species C. *Sci. Rep.* 5:8185. doi: 10.1038/srep08185

Laine, P., Blomqvist, S., Savolainen, C., Andries, K., and Hovi, T. (2006). Alignment of capsid protein VP1 sequences of all human rhinovirus prototype strains, conserved motifs and functional domains. *J. Gen. Virol.* 87, 129–138. doi: 10.1099/vir.0.81137-0

Lamson, D., Renwick, N., Kapoor, V., Liu, Z., Palacios, G., Ju, J., et al. (2006). MassTag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004-2005. *J. Infect. Dis.* 194, 1398–1402. doi: 10.1086/508551

Lau, S. K., Yip, C. C., Tsoi, H. W., Lee, R. A., So, L. Y., Lau, Y. L., et al. (2007). Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic cluster, probably representing a previously undetected HRV species, HRV-C, associated with acute respiratory illness in children. *J. Clin. Virol.* 45, 3655–3664. doi: 10.1128/JCM.01254-07

Ledford, R. M., Collett, M. S., and Pevear, D. C. (2005). Insights into the genetic basis for natural phenotypic resistance of human rhinoviruses to pleconaril. *Antiviral Res.* 68, 135–138. doi: 10.1016/j.antiviral.2005.08.003

Ledford, R. M., Patel, N. R., Demenczuk, T. M., Watanyar, A., Herbertz, T., Collett, M. S., et al. (2004). VP1 sequencing of all human rhinovirus serotypes: insights into genus phylogeny and susceptibility to antiviral capsid-binding compounds. *J. Virol.* 78, 3663–3674. doi: 10.1128/JVI.78.7.3663-3674.2004

Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Slsson, W., et al. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell* 77, 761–771. doi: 10.1016/0092-8674(94)90059-0

McErlean, P., Shackelton, L. A., Lambert, S. B., Nissen, M. D., Sloots, T. P., and Mackay, I. M. (2007). Characterisation of a newly identified human rhinovirus, HRV-QPM, discovered in infants with bronchiolitis. *J. Clin. Virol.* 39, 67–75. doi: 10.1016/j.jcv.2007.03.012

McIntyre, C. L., Knowles, N. J., and Simmonds, P. (2013). Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. *J. Gen. Virol.* 94, 1791–1806. doi: 10.1099/vir.0.053686-0

Miller, E. K., Khuri-Bulos, N., Williams, J. V., Shehabi, A. A., Faouri, S., Al Jundi, I., et al. (2009). Human rhinovirus C associated with wheezing in hospitalised children in the Middle East. *J. Clin. Virol.* 46, 85–89. doi: 10.1016/j.jcv.2009.06.007

Nicholas, K. B., Nicholas, H. B. J., and Deerfield, D. W. (1997). GeneDoc: analysis and visualization of genetic variation, *EMBNEW. NEWS* 4, 14.

Palmenberg, A. C., Rathe, J. A., and Liggett, S. B. (2010). Analysis of the complete genome sequences of human rhinovirus. *J. Allergy Clin. Immunol.* 125, 1190–1199. doi: 10.1016/j.jaci.2010.04.010

Palmenberg, A. C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J. A., et al. (2009). Sequencing and analyses of all known Human Rhinovirus genomes reveal structure and evolution. *Science* 324, 55–59. doi: 10.1126/science.1165557

Paul, A. V., Schultz, A., Pincus, S. E., Oroszlan, S., and Wimmer, E. (1987). Capsid protein VP4 of poliovirus is N-myristoylated. *Proc. Natl. Acad. Sci. U.S.A.* 84, 7827–7831.

Petersen, J. F. W., Cherney, M. M., Liebig, H. D., Skern, T., Kuechler, E., and James, M. N. G. (1999). The structure of the 2A proteinase from a common cold virus,

a proteinase responsible for the shut-off of host-cell protein synthesis. *EMBO J.* 18, 5463–5475. doi: 10.1093/emboj/18.20.5463

Pfister, T., Jones, K. W., and Wimmer, E. (2000). A Cysteine-Rich motif in poliovirus protein 2C ATPASE is involved in RNA replication and binds zinc *in vitro*. *J. Virol.* 74, 334–343. doi: 10.1128/JVI.74.1.334-343.2000

Picornaviridae Study Group (2014). *Rhinovirus C*. Available online at: http://www.picornastudygroup.com/types/enterovirus/hrv-c.htm.

Pilipenko, E. V., Maslova, S. V., Sinyakov, A. N., and Agol, V. I. (1992). Towards identification of cis-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. *Nucleic Acids Res.* 20:1739. doi: 10.1093/nar/20.7.1739

Pond, S. L., and Frost, S. D. (2005). Datamonkey, rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21, 2531–2533. doi: 10.1093/bioinformatics/bti320

Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., et al. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* 317, 145–153. doi: 10.1038/317145a0

Shih, S. R., Chiang, C., Chen, T. C., Wu, C. N., Hsu, J. T., Lee, J. C., et al. (2004). Mutations at KFRDI and VGK domains of enterovirus 71 3C protease affects its RNA binding and proteolytic activities. *J. Biomed. Sci.* 11, 239–248. doi: 10.1007/BF02256567

Simmonds, P., McIntyre, C., Savolainen-Kopra, C., Tapparel, C., Mackay, I. M., and Hovi, T. (2010). Proposals for the classification of human rhinovirus species C into genotypically assigned types. *J. Gen. Virol.* 91, 2409–2419. doi: 10.1099/vir.0.023994-0

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6, molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

van Kuppeveld, F. J. M., Hoenderop, J. G. J., Smeets, R. L. L., Willems, P. H. G. M., Dijkman, H. B. P. M., Galama, J. M. D., et al. (1997). Coxsackievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO J.* 16, 3519–3532. doi: 10.1093/emboj/16.12.3519

Waman, V. P., Kolekar, P. S., Kale, M. M., and Kulkarni-Kale, U. (2014). Population structure and evolution of rhinovirus. *PLoS ONE* 9:e88981. doi: 10.1371/journal.pone.0088981

Wessels, E., Notebaart, R. A., Duijsings, D., Lanke, K., Vergeer, B., Melchers, W. J. G., et al. (2006). Structure-function analysis of the coxsackievirus protein 3A identification of residues important for dimerization, viral RNA replication and transport inhibition. *J. Biol. Chem.* 281, 28232–28243. doi: 10.1074/jbc.M601122200

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Khaw, Chan, Jafar, Othman and Chee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 23 November 2016
doi: 10.3389/fmicb.2016.01886

High Diversity of Genogroup I Picobirnaviruses in Mammals

Patrick C. Y. Woo[1,2,3,4,5†], Jade L. L. Teng[1,2,3,4†], Ru Bai[1†], Annette Y. P. Wong[1], Paolo Martelli[6], Suk-Wai Hui[6], Alan K. L. Tsang[1], Candy C. Y. Lau[1], Syed S. Ahmed[1], Cyril C. Y. Yip[1], Garnet K. Y. Choi[1], Kenneth S. M. Li[1], Carol S. F. Lam[1], Susanna K. P. Lau[1,2,3,4,5] and Kwok-Yung Yuen[1,2,3,4,5*]*

[1] Department of Microbiology, The University of Hong Kong, Hong Kong, China, [2] State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China, [3] Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, China, [4] Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong, China, [5] Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The University of Hong Kong, Hong Kong, China, [6] Ocean Park Corporation, Hong Kong, China

OPEN ACCESS

Edited by:
Akio Adachi,
Tokushima University, Japan

Reviewed by:
Karl William Boehme,
University of Arkansas for Medical Sciences, USA
Yashpal S. Malik,
Indian Veterinary Research Institute, India
Ivana Lojkic,
Croatian Veterinary Institute, Croatia

***Correspondence:**
Patrick C. Y. Woo
pcywoo@hku.hk
Kwok-Yung Yuen
kyyuen@hku.hk

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *08 September 2016*
Accepted: *09 November 2016*
Published: *23 November 2016*

Citation:
Woo PCY, Teng JLL, Bai R, Wong AYP, Martelli P, Hui S-W, Tsang AKL, Lau CCY, Ahmed SS, Yip CCY, Choi GKY, Li KSM, Lam CSF, Lau SKP and Yuen K-Y (2016) High Diversity of Genogroup I Picobirnaviruses in Mammals. Front. Microbiol. 7:1886. doi: 10.3389/fmicb.2016.01886

In a molecular epidemiology study using 791 fecal samples collected from different terrestrial and marine mammals in Hong Kong, genogroup I picobirnaviruses (PBVs) were positive by RT-PCR targeting the partial RdRp gene in specimens from five cattle, six monkeys, 17 horses, nine pigs, one rabbit, one dog, and 12 California sea lions, with 11, 9, 23, 17, 1, 1, and 15 sequence types in the positive specimens from the corresponding animals, respectively. Phylogenetic analysis showed that the PBV sequences from each kind of animal were widely distributed in the whole tree with high diversity, sharing 47.4–89.0% nucleotide identities with other genogroup I PBV strains based on the partial RdRp gene. Nine complete segment 1 (viral loads 1.7×10^4 to 5.9×10^6/ml) and 15 segment 2 (viral loads 4.1×10^3 to 1.3×10^6/ml) of otarine PBVs from fecal samples serially collected from California sea lions were sequenced. In the two phylogenetic trees constructed using ORF2 and ORF3 of segment 1, the nine segment 1 sequences were clustered into four distinct clades (C1–C4). In the tree constructed using RdRp gene of segment 2, the 15 segment 2 sequences were clustered into nine distinct clades (R1–R9). In four sea lions, PBVs were detected in two different years, with the same segment 1 clade (C3) present in two consecutive years from one sea lion and different clades present in different years from three sea lions. A high diversity of PBVs was observed in a variety of terrestrial and marine mammals. Multiple sequence types with significant differences, representing multiple strains of PBV, were present in the majority of PBV-positive samples from different kinds of animals.

Keywords: diversity, picobirnaviruses, mammals, genogroup I, sea lion

INTRODUCTION

Picobirnaviruses (PBVs) are small non-enveloped bisegmented double-stranded RNA viruses found in human and a wide variety of mammals and birds. Since its first discovery in fecal samples of humans and rats in Pereira et al. (1988a,b), PBVs have been reported in a variety of other terrestrial mammals, birds and environmental water samples (Fregolente et al., 2009; Ghosh et al., 2009; Symonds et al., 2009; Martinez et al., 2010; Ganesh et al., 2011a; Malik et al., 2011; Wang et al., 2012; Bodewes et al., 2013; Gillman et al., 2013; Ng et al., 2014; Ribeiro et al., 2014; Zhang et al., 2014). In 2012, we reported the discovery of a PBV, named otarine PBV (Ot-PBV), in a

California sea lion (*Zalophus californianus*) in Hong Kong, which was the first PBV reported in a marine mammal (Woo et al., 2012). Recently, we have also described the first discovery and a diversity of PBVs in dromedary camels from the Middle East (Woo et al., 2014b).

The genome of PBV consists of two segments named segment 1 and segment 2. Segment 1 contains the capsid gene and another open reading frame which encodes for a putative protein of unknown function, whereas segment 2 contains the RNA-dependent RNA polymerase (RdRp) gene (Wakuda et al., 2005; Woo et al., 2012). By sequence and phylogenetic analyses, PBVs are classified into genogroups I and II based on the RdRp gene sequence (Rosen et al., 2000; Wakuda et al., 2005). Recently, it has also been reported that novel genogroups of PBVs have been detected in human and environmental samples (Smits et al., 2014; Zhang et al., 2015). As of December 31 2015, 931 PBV nucleotide sequences have been submitted to GenBank. However, only nine are complete/near-complete segment 1 sequences and 21 are complete/near-complete segment 2 sequences. Among these nine complete/near-complete segment 1 and 21 complete/near-complete segment 2 sequences, 18 are genogroup I sequences.

Our recent study on dromedary camel PBVs (Woo et al., 2014b) and our preliminary analysis using the limited PBV nucleotide sequences in GenBank (data not shown) revealed that different genogroup I PBVs could be present in different animals of the same species. Since a high diversity of genogroup I PBVs may exist in different animals, we performed a molecular epidemiology study using fecal samples collected from different terrestrial and marine mammals in Hong Kong. In addition, we studied the evolution of genogroup I PBVs in California sea lions by serially collecting their fecal samples for 6 years and sequenced and analyzed the complete segments 1 and 2 of the PBV-positive samples.

MATERIALS AND METHODS

Terrestrial and Marine Mammal Surveillance and Sample Collection

This study was performed in strict accordance with local ordinance and the recommendations by the Committee on the Use of Live Animals in Teaching and Research (CULATR) at The University of Hong Kong. All specimens of bats, monkeys, cats, and dogs were collected with the assistance of the Department of Agriculture, Fisheries and Conservation, Hong Kong Special Administrative Region (HKSAR); those of pigs and cattle were collected with the assistance of the Department of Food, Environmental and Hygiene, HKSAR, from various locations in HKSAR; and those of horses were collected with the assistance of the Hong Kong Jockey Club. All specimens of rabbits were collected from live food animal markets in Guangzhou, China, in October 2007. Rectal swabs were collected using procedures described previously (Lau et al., 2005). All fecal samples of marine mammals, including Indo-Pacific bottlenose dolphins, California sea lions and harbor seals, were collected by veterinary surgeons of the Ocean Park in HKSAR (Woo et al., 2014a). A total of 791 samples collected over a 75-month period (October 2007 to December 2013) from 157 bats, 52 monkeys, 100 pigs, 58 cats, 58 dogs, 50 cattle, 106 rabbits, 95 horses, 46 Indo-Pacific bottlenose dolphins, 54 California sea lions, and 15 harbor seals were tested (**Table 1**).

TABLE 1 | Prevalence of PBVs in mammals by RT-PCR targeting a partial fragment of RdRp gene.

Animal type	Scientific name	Number of samples obtained	Number (%) of samples positive for genogroup I PBVs	Health status
Terrestrial mammals				
Cattle	NA	50	6 (12.0)	Deceased
Monkey*	NA	52	6 (11.5)	Healthy
Horse†	NA	95	17 (18.0)	Sick with fever
Pig	NA	100	9 (9.0)	Deceased
Rabbit	*Oryctolagus cuniculus*	106	1 (0.9)	Healthy
Dog	*Canis familiaris*	58	1 (2.0)	Healthy
Cat	*Feline domestica*	58	0 (0)	Healthy
Bat	*Hipposideros pomona*	103	0 (0)	Healthy
	Rhinolophus affinis	14	0 (0)	Healthy
	Rhinolophus sinicus	40	0 (0)	Healthy
Marine mammals				
California sea lion	*Zalophus californianus*	54	12 (22.0)	Healthy
Indo-Pacific bottlenose dolphin	*Tursiops aduncas*	46	0 (0)	Healthy
Harbor seal	*Phoca vitulina*	15	0 (0)	Healthy

**Most monkeys found in our locality are hybrids of Macaca mulatta and Macaca fascicularis. †All horses included in this study were thoroughbred racehorses stabled at the Hong Kong Jockey Club.*
NA, not available.

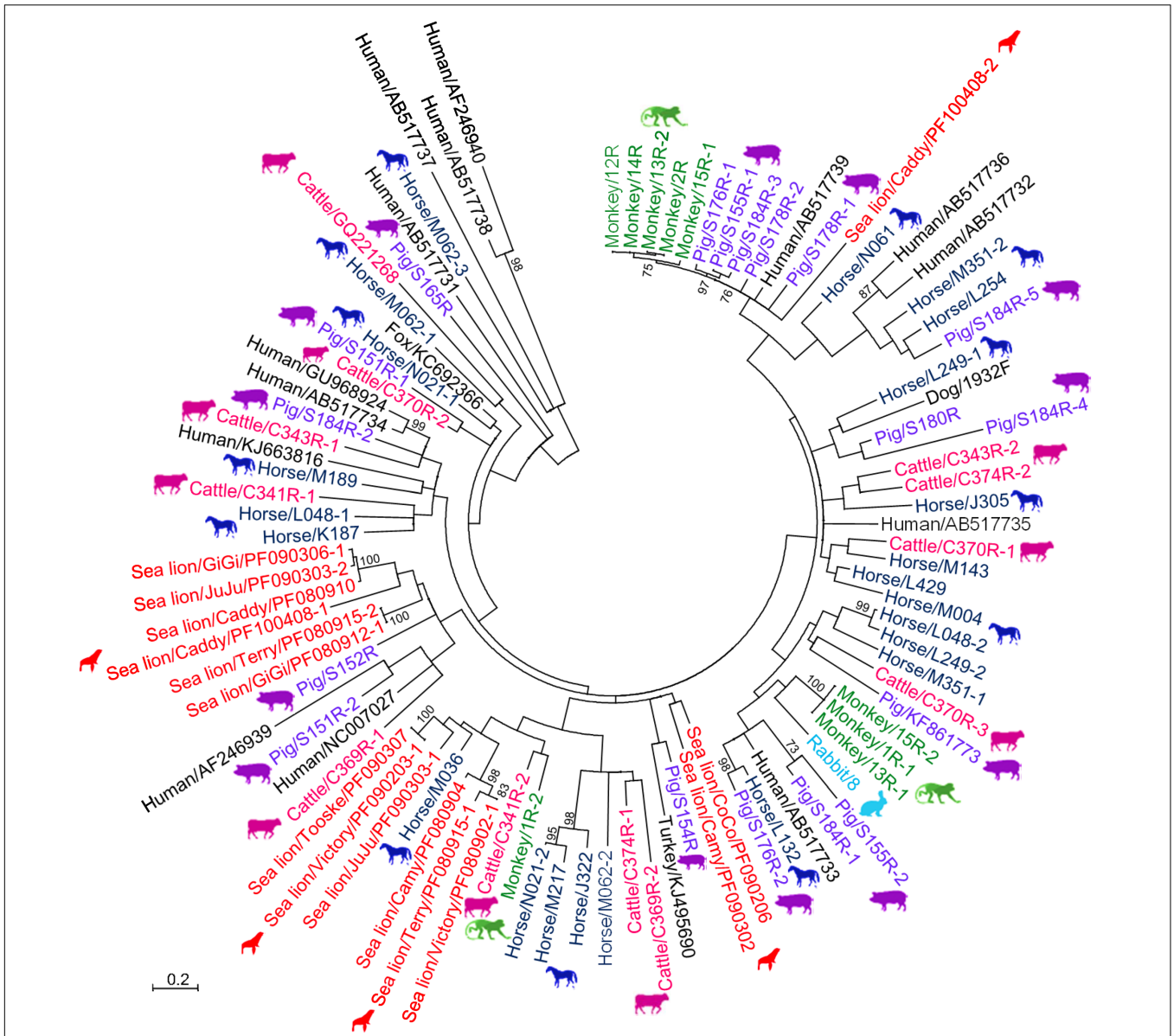


FIGURE 1 | Phylogenetic analysis of partial RdRp gene of genogroup I PBVs discovered in the present study. The tree was constructed using the maximum-likelihood method and the optimal substitution model of T92+G+I, and rooted with genogroup II human strains, shown in italics. One hundred and ninety one nucleotide positions were included in the analysis. Bootstrap values below 70% are not shown. The scale bar indicated the number of nucleotide substitutions per site. All PBV strains discovered in this study are colored, with those detected in the same host highlighted in the same color. If more than one sequence type was found in the same sample, each sequence type was numbered in the order of identification (e.g., Monkey/13R-1, Monkey/13R-2, and Monkey/13R-3 indicated that there were three sequence types found in the same monkey sample 13R). All the accession numbers are given as cited in GenBank.

RNA Extraction

Viral RNA was extracted from rectal and cloacal swabs and fecal samples using EZ1 Virus Mini Kit v2.0 (Qiagen, Germany). RNA was eluted in 60 μl of AVE buffer (Qiagen, Germany) and about 200 ng of RNA was used as template for RT-PCR.

RT-PCR for PBVs and DNA Sequencing

Genogroup I PBV screening was performed by PCR amplification of a 205-bp fragment of the RdRp gene of genogroup I PBVs using conserved primers (5′-CAAARTTYGACCARCACTT-3′ and 5′-TCRTCDGCRTTGGTACCACC-3′) designed by multiple alignments of the available RdRp genes of PBVs. Reverse transcription was performed using the SuperScript III kit (Invitrogen, USA) and the reaction mixture (10 μl) contained RNA, first-strand buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$), 5 mM DTT, 50 ng random hexamers, 500 μM of each dNTPs and 100 U Superscript III reverse transcriptase. The mixtures were incubated at 25°C for 5 min, followed by 50°C for 60 min and 70°C for 15 min. The PCR mixture (25 μl) contained

TABLE 2 | Genomic features and coding potential of otarine PBVs detected in California sea lions in this study.

California sea lion	Sampling year	Segment 1						Segment 2					
		Clade	ORFs	Location (nt)	Length (nt)	Length (aa)	Frame	Clade	ORFs	Location (nt)	Length (nt)	Length (aa)	Frame
Victory	2008							R3	RdRp	45–1643	1599	532	3
	2009	C3	ORF1	45–230	186	61	3	R4	RdRp	46–1635	1590	529	1
			ORF2	136–510	375	124	1						
			ORF3	537–2126	1590	529	3						
	2009	C4	ORF2	164–835	672	223	2						
			ORF3	832–2502	1671	556	1						
GiGi	2008	C3	ORF1	45–230	186	61	3	R2	Novel ORF	39–185	147	48	3
			ORF2	136–510	375	124	1		RdRp	167–1783	1617	538	2
			ORF3	537–2126	1590	529	3						
	2009	C2	ORF2	170–847	678	225	2	R1	Novel ORF	46–210	165	54	1
			ORF3	850–2487	1638	545	1		RdRp	251–1846	1596	531	2
	2009	C3	ORF1	45–230	186	61	3						
			ORF2	136–510	375	124	1						
			ORF3	537–2126	1590	529	3						
Caddy	2008	C4	ORF2	165–836	672	223	3	R1	Novel ORF	46–210	165	54	1
			ORF3	833–2503	1671	556	2		RdRp	251–1846	1596	531	2
	2010							R8	Novel ORF	47–262	216	71	2
									RdRp	285–1874	1590	529	3
	2010							R9	RdRp	140–1759	1620	539	2
Camy	2008							R3	RdRp	46–1644	1599	532	1
	2009							R7	RdRp	168–1772	1605	534	3
JuJu	2009	C2	ORF2	170–847	678	225	2	R1	Novel ORF	48–212	165	54	3
			ORF3	850–2487	1638	545	1		RdRp	253–1848	1596	531	1
	2009							R5	RdRp	46–1638	1593	530	1
Terry	2008	C1	ORF1	55–198	144	47	1	R2	Novel ORF	39–185	147	48	3
			ORF2	104–688	585	194	2		RdRp	167–1783	1617	538	2
			ORF3	703–2430	1728	575							
	2008						1	R3	RdRp	46–1644	1599	532	1
Tooske	2009	C4	ORF2	165–836	672	223	3	R4	RdRp	46–1635	1590	529	1
			ORF3	833–2503	1671	556	2						
CoCo	2009							R6	RdRp	296–1897	1602	533	2

cDNA, PCR buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 2 mM $MgCl_2$), 200 μM of each dNTPs and 1.0 U *Taq* polymerase (Applied Biosystems, USA). The mixtures were amplified in 60 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1 min and a final extension at 72°C for 10 min in an automated thermal cycler (Applied Biosystems, USA).

All PCR products were gel-purified using the QIAquick gel extraction kit (Qiagen, Germany). Both strands of the PCR products were sequenced twice with an ABI Prism 3730xl DNA Analyzer (Applied Biosystems, USA), using the two PCR primers. As multiple nucleotide peaks were observed in most sequencing results, it was suggested that more than one type of PBV were present in each sample, thus the purified PCR products were cloned into the pCR-II-TOPO TA cloning vector (Invitrogen, USA) according to manufacturer's instructions. Both strands of 10 clones for each sample were sequenced, using primers 5′-TAATACGACTCACTATAGGG-3′ and 5′-CGGCTCGTATGTTGTGTGGA-3′. The sequences of the clones were compared with known sequences of the RdRp of PBVs in the GenBank database.

Complete Segments 1 and 2 Sequencing of Genogroup I Otarine PBVs

Nine complete segments 1 and 15 complete segment 2 of otarine PBVs were amplified and sequenced using published strategies for double-stranded RNA viruses (Attoui et al., 2000), using RNA extracted from the original specimens of sea lions positive for PBV as template. Viral RNA was extracted using the EZ1 virus mini kit (Qiagen, Germany). Adaptor primer, with 3′ NH_2 blocking group, was ligated to the 3′ termini of the viral RNA and subjected to reverse transcription using complementary primer. After RNA hydrolysis, reannealing and end-filling, single-primer amplification of viral genomic segments was performed using complementary primer and genome specific primers. The 5′ and 3′ ends of the viral genomes were confirmed by rapid amplification of cDNA ends using

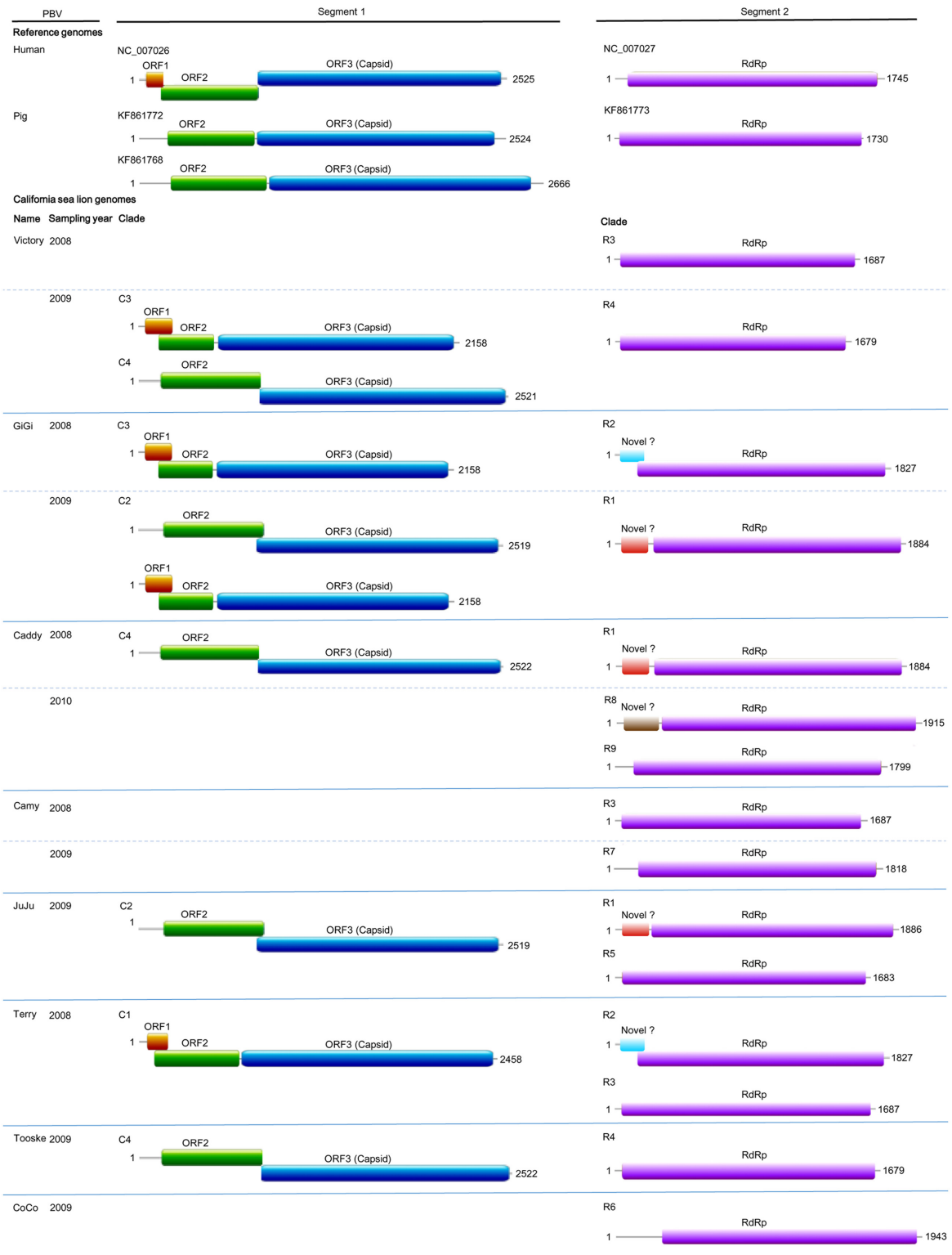


FIGURE 2 | Genome organization of otarine PBVs and the two representative genogroup I PBVs. The name of each California sea lion, year of sample collection and clade (C1–C4 and R1–R9) for each otarine PBV are indicated. For the nine segment 1 sequences, the capsid protein is represented by a blue box. Upstream to the capsid protein, ORF1 and ORF2 are represented by orange and green boxes respectively. For the 15 segment 2 sequences, the RdRp is represented by a purple box. The three different kinds of previously undescribed ORF upstream to the RdRp in six of the 15 segment 2 are represented by boxes highlighted with different colors.

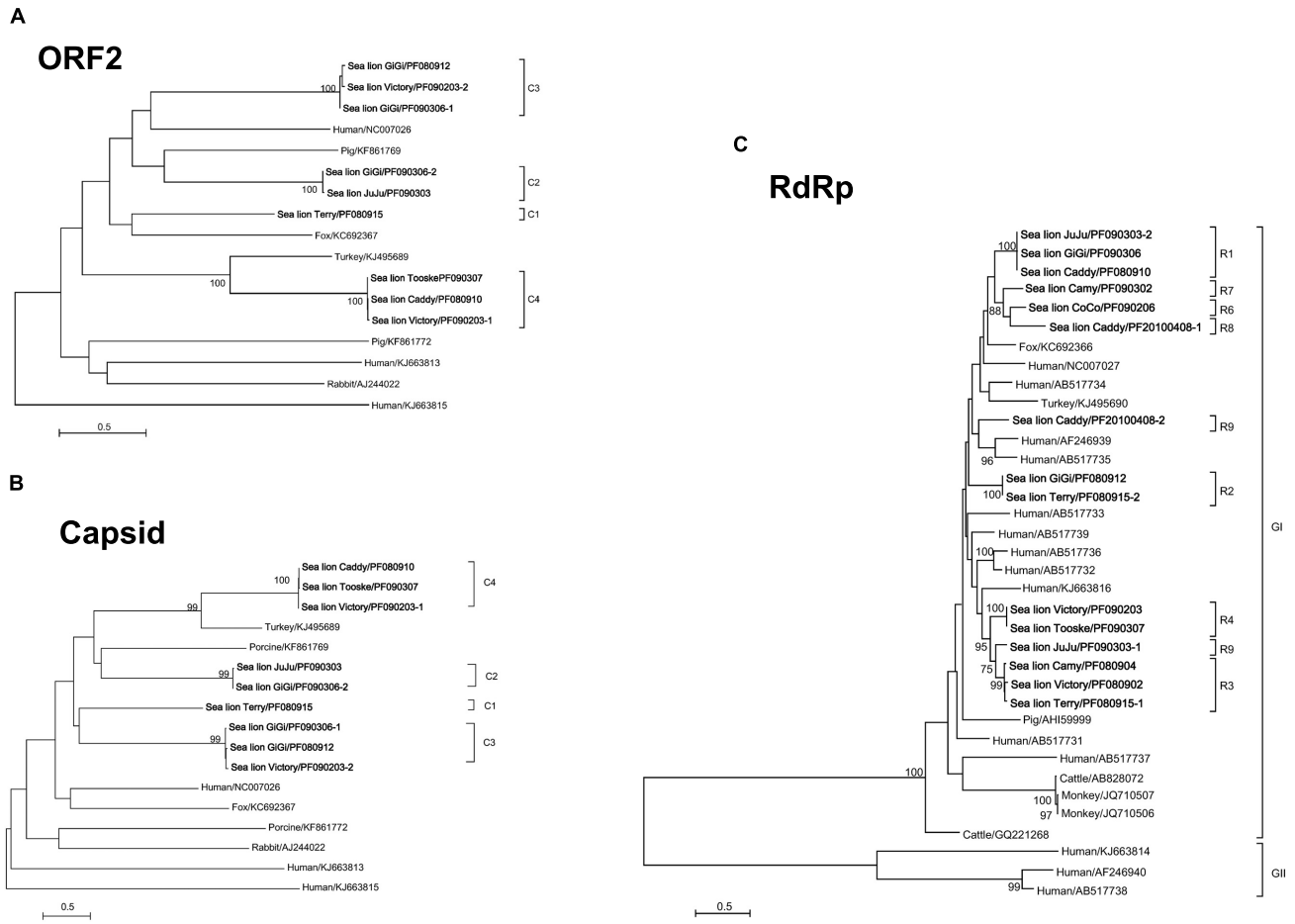


FIGURE 3 | Phylogenetic analyses of ORF2 (A), capsid **(B)**, and RdRp proteins of otarine PBVs **(C)**. The trees are constructed using maximum likelihood method and the optimal substitution models of LG+G+F for ORF2 and LG+G for capsid and RdRp proteins. Two hundred and seventeen, 545 and 475 amino acid positions in ORF2, capsid and RdRp proteins were included respectively in the analyses. The scale bar indicated the number of nucleotide substitutions per site. Bootstrap values below 70% are not shown. The otarine PBVs characterized in this study are bolded. If more than one sequence type was found in the same sample, each sequence type was numbered in the order of identification (e.g., Victory/PF090203-1 and Victory/PF090203-2 indicated that there were two sequence types found in the same California sea lion sample Victory/PF090203). The clade of each otarine PBV is also indicated. All the accession numbers are given as cited in GenBank. GI and GII represent genogroup I and genogroup II PBVs respectively.

the 5′/3′ RACE kit (Roche, Germany). The PCR products were gel purified and sequenced using an ABI Prism 3700 DNA analyzer (Applied Biosystems, USA). Sequences were assembled and manually edited to produce the final sequences of the viral genomes.

Genome Analysis

The nucleotide sequences of the genomes and the deduced amino acid sequences of the ORFs were compared to those of other PBVs. Novel genes were further predicted by FGENESV (SoftBerry, Inc.[1]), a trained pattern/Markov chain-based viral gene prediction program. Phylogenetic tree construction was performed using the maximum likelihood method and MEGA7 (Kumar et al., 2016), with bootstrap values being calculated from 1,000 trees. The optimal substitution model for each ORF was selected by MEGA7. Protein domain, family and functional site analyses were performed using ScanProsite (De Castro et al., 2006). Transmembrane and coiled-coil domains were predicted by TMHMM Server v 2.0 (Krogh et al., 2001) and COILS (Lupas et al., 1991) respectively.

Quantitative RT-PCR

For real-time quantitative PCR assays, cDNA were amplified in SYBR Green I fluorescence reactions (Roche, Germany). Briefly, 10 μl of reaction mixtures containing 1 μl cDNA, 10 μl FastStart DNA master SYBR green I mix reagent (Roche) and 5 mM each of forward and reverse specific primers were thermal-cycled at 95°C for 10 min followed by 45 cycles of 95°C for 10 s, 60°C for 10 s and 72°C for 20 s using a Roche LightCycler 96 real time PCR system (Roche, Germany). Specific primers for each PBV segment were designed based on the sequences of all the segment 1 and segment 2 detected in the positive samples (**Table 3**).

[1]www.softberry.com/

Plasmids with the corresponding target sequences were used for generating the standard curve. At the end of the assay, PCR products were subjected to a melting curve analysis (65–95°C, 0.1°C/s) to confirm the specificity of the assay.

RESULTS

Detection of Diverse Genogroup I PBVs in Animals

A total of 791 fecal specimens from 676 terrestrial mammals and 115 marine mammals were obtained (**Table 1**). RT-PCR for a 205-bp fragment in the RdRp gene of genogroup I PBVs was positive in specimens from six cattle, six monkeys, 17 horses, nine pigs, one rabbit, one dog, and 12 sea lions. Marked nucleotide polymorphisms were observed in most of the RdRp sequences, suggesting the possible existence of multiple strains in the same specimen. Therefore, the PCR products were cloned and 10 clones from each specimen were sequenced. Multiple sequence types were confirmed to be present in most samples. Sequence analysis of these clones revealed that there were 11, 9, 23, 17, 1, 1, and 15 sequence types in the positive specimens from the five cattle, six monkeys, 17 horses, nine pigs, one rabbit, one dog, and 12 sea lions respectively, and 47.4–89.0% nucleotide identities were observed between these clones and the corresponding sequences of other genogroup I PBV strains available in the GenBank database (**Figure 1**). The PBV sequences from each kind of animal were widely distributed in the whole phylogenetic tree with high diversity (**Figure 1**). No PBV was detected in the specimens obtained from the 58 cats, 157 bats, 46 Indo-Pacific bottlenose dolphins, and 15 harbor seals (**Table 1**).

Otarine PBVs Complete Segments 1 and 2 Sequence Analysis

Nine complete segment 1 of otarine PBVs from sea lions were sequenced and assembled (**Table 2**; **Figure 2**). These segment 1 sequences ranged from 2,158 to 2,522 bases in length with overall G+C contents of 41.1–46.0%. The 5′ non-coding regions (44–169 bases) were AU-rich (G+C contents of 22.7–37.9%) with five conserved bases, GUAAA, located at the 5′ end. A predicted highly stable stem loop structure found in other known PBVs was observed in five segment 1 sequences as a result of the pairing of 5′-GUAAA-3′ and 5′-UUUAC-3′ in the 5′ non-coding region (Nates et al., 2011). The 3′ non-coding regions (19–32 bases) contained G+C contents ranging from 53.1 to 71.4% and end with 3–4 conserved bases (CTC, CTTC, or CACC). All the nine segment 1 sequences possess one long ORF (1,590–1,728 bp) encoding the capsid protein of 529–575 amino acids. These capsid proteins shared low (19.3–37.2%) amino acid identities with those of other PBV strains, being most closely related to turkey PBV TK/MN/2011 (GenBank number KJ495689), fox PBV F5-1 (GenBank number KC692367) and human PBV Hy005102 (GenBank number NC_007026). Upstream to the ORF for the capsid protein, there were one to two short ORFs in the nine segment 1, consistent with the organization of the segment 1 in other known PBVs (Wakuda et al., 2005; Bodewes et al., 2013; Banyai et al., 2014; Verma et al., 2015). The protein encoded by the ORF2 of segment 1 from the nine otarine PBVs possessed different numbers of repetitions of the same motif, ExxRxNxxxE, that was also observed in the corresponding protein in other known PBVs (Da Costa et al., 2011).

Fifteen complete segment 2 from the otarine PBVs were sequenced and assembled (**Table 2**; **Figure 2**). These segment 2 sequences ranged from 1,679 to 1,943 bases in length with overall G+C contents of 39.5–48.1%. The 5′ non-coding regions (38–295 bases) were also AU-rich (G+C contents of 18.4–35.9%) with the same conserved bases, GUAAA, located at the 5′ end. The stable stem loop structure observed in segment 1 was also observed in the 5′ non-coding regions of 13 segment 2 sequences. The 3′ non-coding regions (38–46 bases) have G+C contents ranging from 31.6 to 56.1% and end with four conserved bases (CUGC) in most of the genomes. All the 15 segment 2 sequences possess one long ORF (1,590–1,620 bp) encoding the RdRp of 529–539 amino acids. These RdRp shared 44.5–70.6% amino acid identities with those of other genogroup I PBV strains, being most closely related to fox PBV F5-1 (GenBank number KC692366), human PBV 1-CHN-97 (GenBank number AF246939), human PBV GPBV6C1 (GenBank number AB517731), human PBV HuPBV-E-CDC16 (GenBank number KJ663816) and human PBV VS10 (GenBank number GU968924). They possess three conserved motifs (D-T/S-D, SG-T, GDD) commonly found in the RdRp sequences of other dsRNA viruses. Conserved cysteine and proline residues present in other genogroup I PBVs were also observed in all 15 segment 2 sequences. In contrast to the segment 2 sequences of other known PBVs, six of our 15 sequenced segment 2 possess a previously undescribed ORF 48–71 amino acids upstream to the RdRp ORF. Multiple alignments of the sequences of these six ORFs showed that they formed three groups that were not homologous to each other, with three ORFs belonging to the first group, two ORFs to the second group and one ORF to the third group (**Figure 2**). In the first group, the three ORFs showed 97.6–99.4% nucleotide identities among each other. In the second group, the two ORFs showed 100% nucleotide identity. Sequence analysis of these ORFs did not reveal any significant sequence homology to other proteins in the GenBank database. Protein sequence analysis also did not reveal any significant matches to other known protein domains, families or functional sites in the PROSITE database. Moreover, there are no transmembrane and coiled-coil domains predicted by sequence analyses in these protein sequences.

Phylogenetic Analysis of Otarine PBVs Complete Segments 1 and 2 Sequences

In the phylogenetic trees constructed using ORF2 (**Figure 3A**) and ORF3 (capsid protein) (**Figure 3B**), the nine sequenced segment 1 of the otarine PBVs were clustered into four distinct clades (C1–C4) (**Figures 3A,B**). The sequence types that belonged to each clade were identical for both trees, suggesting that there was no recombination between different PBVs. In the phylogenetic tree constructed using the RdRp gene, the 15 sequenced segment 2 of the otarine PBVs were clustered into nine distinct clades (R1–R9) (**Figure 3C**). Although, R1, R6, R7, and

TABLE 3 | Quantitative RT-PCR assays of the California sea lions positive for PBV.

Name of California sea lion	Sample number	Segment 1/2 – sequence type	RNA concentration (copies/ml)	Sampling dates (YYYYMMDD)	Forward primer (5′–3′)	Reverse primer (5′–3′)
Victory						
	PF080902	Segment 2	2.4×10^4	20080911	CCATGACTCAAGTCCTTACTAGT	TTTAGTGATCCGCTGGTCGACGGCA
	PF090203	Segments 1-1	1.4×10^6	20090209	GAAGATGGTTATACCATTAGTGGTTA	AACTGTGAAGGAATTTTGAATCTG
	PF090203	Segments 1-2	2.2×10^4	20090209	GAAGATTGTATGACCATGGGCGGTGA	AATGAACGTAACATTCATAATCTG
	PF090203	Segment 2	1.9×10^4	20090209	TCGCTTCCGATATGACTCAA	ACATCTTAGTGATACGCTGG
Camy						
	PF080904	Segment 2	9.8×10^4	20080911	CCATGACTCAAGTCCTTACTAGT	TTTAGTGATCCGCTGGTCGACGGCA
	PF090302	Segment 2	5.4×10^4	20090304	ACTAGCCACGGTCTGGAGAT	ACACACACCATCTTGCCGAT
Caddy						
	PF080910	Segment 1	1.7×10^4	20080912	GAAGATGGTTATACCATTAGTGGTTA	AACTGTGAAGGAATTTTGAATCTG
	PF080910	Segment 2	4.1×10^3	20080912	GGTTTGGACGCATATCAAGA	ACATTCTAGTTATAGCTCGA
	PF100408	Segment 2-1	7.3×10^4	20100425	AGATTGCTGTACCAGGAGCG	GCGCTACCATCTTAGGACCC
	PF100408	Segment 2-2	6.9×10^5	20100425	TGGGTAGTGGTTCAGGAGGT	CGGTTATACCGGGGTACGTG
CoCo						
	PF090206	Segment 2	7.4×10^5	20090217	GGTTGTGTTACCAGGAGCGA	CTTAATTGCCGCAGAGCGAC
GiGi						
	PF080912	Segment 1	9.7×10^4	20080912	GAAGATTGTATGACCATGGGCGGTGA	AATGAACGTAACATTCATAATCTG
	PF080912	Segment 2	3.3×10^5	20080912	GGAGATGACGTGGCACAGG	ATAGACGAGTTATCGCTTTA
	PF090306	Segment 1-1	5.9×10^6	20090310	GAAGATTGTATGACCATGGGCGGTGA	AATGAACGTAACATTCATAATCTG
	PF090306	Segment 1-2	2.2×10^6	20090310	GAAGATTTTGCAATTATGAGCGGTGA	AATAGTACGAGTGTTTGCAATCTG
	PF090306	Segment 2	1.3×10^6	20090310	GGTTTGGACGCATATCAAGA	ACATTCTAGTTATAGCTCGA
Terry						
	PF080915	Segment 1	4.6×10^6	20080912	GAAGATATGGCTATTATGAGTGGTGA	CCAGATTGTAGCATTTTCGATCTG
	PF080915	Segment 2-1	7.8×10^5	20080912	CCATGACTCAAGTCCTTACTAGT	TTTAGTGATCCGCTGGTCGACGGCA
	PF080915	Segment 2-2	1.2×10^5	20080912	GGAGATGACGTGGCACAGG	ATAGACGAGTTATCGCTTTA
JuJu						
	PF090303	Segment 1	1.0×10^6	20090304	GAAGATTTTGCAATTATGAGCGGTGA	AATAGTACGAGTGTTTGCAATCTG
	PF090303	Segment 2-1	4.2×10^5	20090304	GCATCTACTATGACCCAGGTT	ACATCTTAGTGATCCGCCGA
	PF090303	Segment 2-2	6.2×10^4	20090304	GGTTTGGACGCATATCAAGA	ACATTCTAGTTATAGCTCGA
Tooske						
	PF090307	Segment 1	2.1×10^5	20090318	GAAGATGGTTATACCATTAGTGGTTA	AACTGTGAAGGAATTTTGAATCTG
	PF090307	Segment 2	2.1×10^4	20090318	TCGCTTCCGATATGACTCAA	ACATCTTAGTGATACGCTGG

TABLE 4 | Prevalence of PBVs identified in serial fecal samples of California sea lions.

Name of California sea lion	Detection of PBV (clade* for each otarine PBV is indicated in bracket)			
	2008	2009	2010	2013
Beamer	–	–	–	NA
Victory	+ (R3)	+ (C3, C4, R4)	–	–
BoyBoy	–	–	–	NA
Camy	+ (R3)	+ (R7)	–	NA
Wahoo	–	–	–	NA
Baja	–	–	NA	NA
Bobo	–	–	–	NA
Caddy	+ (C4, R1)	–	+ (R8, R9)	NA
Coco	–	+ (R6)	–	–
GiGi	+ (C3, R2)	+ (C2, C3, R1)	NA	NA
Julius	–	–	NA	NA
Jumanji	–	–	NA	–
Terry	+ (C1, R2, R3)	–	NA	NA
Beeper	–	–	–	–
JuJu	–	+ (C2, R1, R5)	NA	–
Jellybean	NA	–	NA	–
Tooske	NA	+ (C4, R4)	–	–
Julie	NA	–	NA	–

**C1–C4 and R1–R9.*
+, positive for PBVs; –, negative for PBVs.
NA, samples not available.

R8 as well as R3, R4, and R5 seemed to be further clustered in the tree, there were >14% amino acid difference between any two clades.

Quantitative RT-PCR

Quantitative RT-PCR showed that the amounts of otarine PBV RNA ranged from 1.7×10^4 to 5.9×10^6 (for segment 1) and 4.1×10^3 to 1.3×10^6 (for segment 2) copies/ml in the fecal samples (**Table 3**).

Longitudinal Detection of Otarine PBVs and Genome Evolution

Fecal samples were serially obtained from 18 sea lions over 6 years (**Table 4**). In the 12 samples that were positive for PBV, segment 2 could be detected and sequenced in all 12 samples, but segment 1 could only be detected and sequenced in seven samples due to difficulties in designing PCR primers as a result of the limited number of segment 1 PBV sequences available in GenBank. Overall, samples positive for PBV were collected mainly in 2008 and 2009. Among these positive results, most clades of segment 1 and segment 2 were observed in samples from either 2008 or 2009, while only segment 1 clades C3 and C4 and segment 2 clade R1 were present in samples from both 2008 and 2009. More than one clade of segment 2 was present in three of the 12 samples and more than one clade of segment 1 was present in two of the seven segment 1 positive samples. In four sea lions (Victory, Camy, Caddy, and GiGi), PBVs were detected in samples collected from two different years (**Table 4**). Among these four sea lions, the same segment 1 clade (C3) was present in two consecutive years in one sea lion (GiGi), with a total of 18 nucleotide changes in the capsid coding region of the same clade (C3) detected between 2008 and 2009. In the other three sea lions (Victory, Camy and Caddy), different clades were present in the fecal samples obtained in different years.

Nucleotide Sequence Accession Numbers

The genome sequences of otarine PBVs obtained from the present study were deposited in GenBank with accession numbers KU729746–KU729769.

DISCUSSION

A high diversity of PBVs was observed in a variety of terrestrial and marine mammals. Despite the relatively high evolutionary rate of RNA viruses, those that infect a specific host usually fall into several discrete viral species; for example, human coronaviruses that infect human include four distinct species, namely OC43 (Vabret et al., 2003), 229E (Yeager et al., 1992), NL63 (Hofmann et al., 2005), and HKU1 (Woo et al., 2005). Viruses infecting a specific host do not form a "continuous" spectrum. However, when we tried to perform phylogenetic analysis on short fragments of PBVs amplified from human samples downloaded from the GenBank, it was noted that these human PBVs formed a spectrum covering the whole phylogenetic tree (data not shown). In the present study on PBVs of different animal hosts, a similar phenomenon was observed. PBVs from horses, pigs, and cattle were widely distributed in the whole phylogenetic tree, and PBVs from sea lions and monkeys also

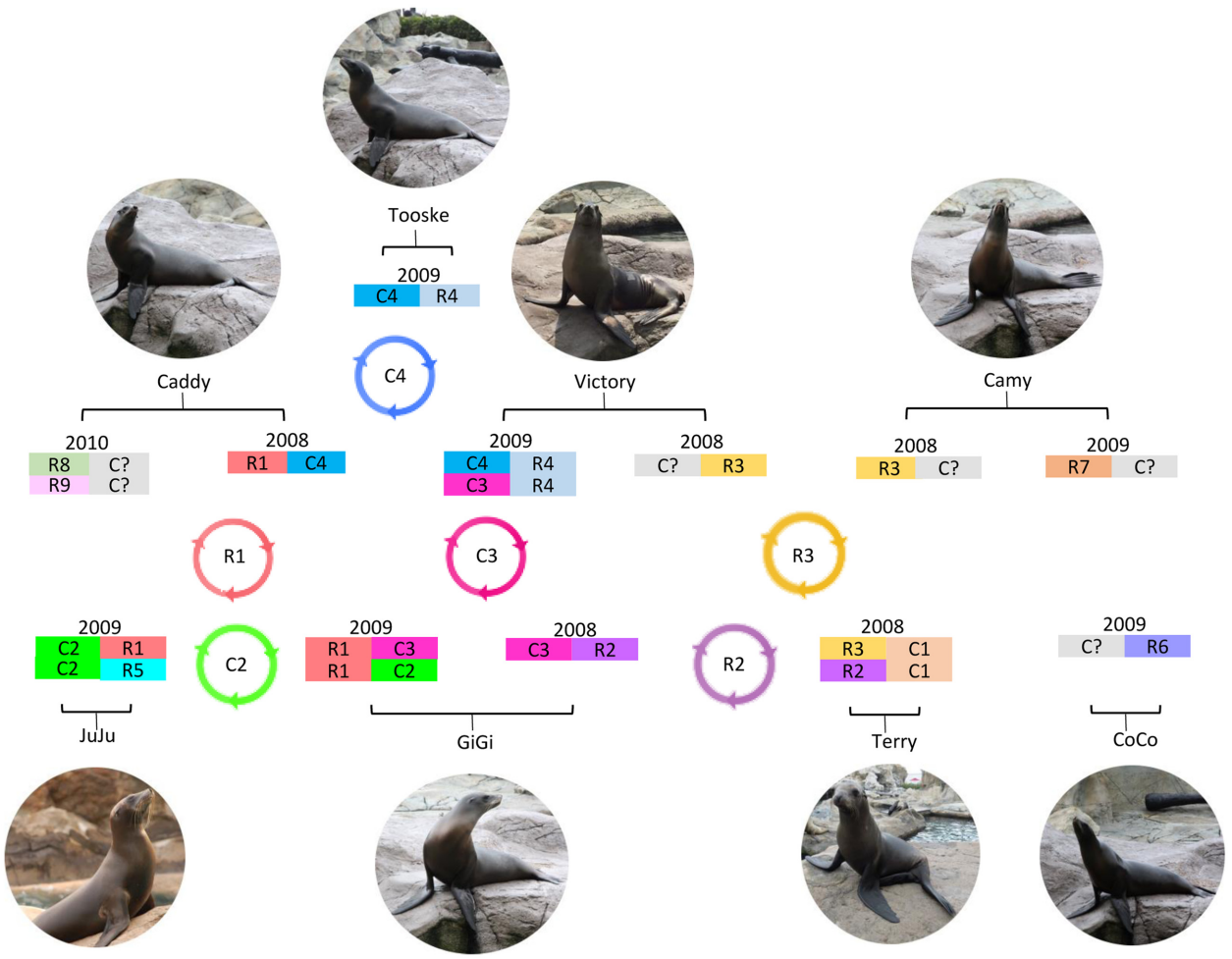


FIGURE 4 | Circulation and evolution model of otarine PBVs in California sea lions. Possible otarine PBVs detected in the serial fecal samples of California sea lions collected in different years are shown. Possible reassortment of segments 1 and 2 of different clades (C1–C4 and R1–R9) circulating in the California sea lions are also indicated.

showed high diversity (**Figure 1**). In addition no two PBVs were detected to be the same among the 52 positive samples, except for that found in two pairs of sea lions, which possessed identical sequences even in the short fragment of 205 bases used for screening. The reason for this phenomenon remains unclear, although it may be partly due to the high mutation rate of PBVs, as quasispecies are frequently found in PBV sequences. Unfortunately, no PBV has been isolated in all studies so far. Therefore, further confirmation by repeated passage of a PBV strain and sequencing, which will determine its mutation rate more accurately, is not possible.

Multiple strains of PBV were present in the majority of PBV-positive samples from different kinds of animals. In the literature, the presence of multiple strains of PBV has only been described in fecal samples collected from human (van Leeuwen et al., 2010; Ganesh et al., 2011b), pigs (Bányai et al., 2008; Chen et al., 2014), chickens (Ribeiro et al., 2014), monkeys (Wang et al., 2012) and buffalos (Malik et al., 2014), despite the identification of PBVs in 24 different animals. In this study, multiple strains of PBV were observed in the same fecal samples of cattle, monkeys, horses, pigs, and sea lions. This phenomenon was also confirmed by sequencing the complete segments 1 and 2 of otarine PBVs directly from the fecal samples of sea lions (**Figure 2**; **Table 4**). Among the 12 samples that showed positive results, more than one clade of segment 2 were present in three of the 12 samples and more than one clade of segment 1 were present in two of the seven segment 1 positive samples (**Table 4**). It is of note that due to the low number of complete segments 1 and 2 sequences of PBV in GenBank, some of the segments in the fecal samples of the sea lions in the present study could not be amplified and sequenced. In fact, only nine complete/near-complete PBV segment 1 sequences are available in GenBank, making sequencing of segment 1 particularly difficult using the genome walking approach. Despite these technical difficulties, more than one segment 1 and/or more than one segment 2 were observed in at least five fecal samples of the sea lions in this study. The presence of more than one segment 1 and more than one segment 2 in the same sample is rare in other segmented RNA

viruses. This phenomenon makes it difficult to ascertain which segment 1 corresponds to which segment 2 in individual PBV genomes from a specific fecal sample.

In the PBV genomes sequenced in this study, only the two ORFs that encode for the capsid protein and RdRp showed significant homologies to the corresponding ORFs encoding the same proteins in other PBVs. In addition to these two ORFs, an ORF1 and an ORF2 were found upstream to the ORF that encodes for the capsid protein in segment 1 of some PBVs. Bioinformatics analysis showed that the deduced amino acid sequences of ORF1 and ORF2 possess no significant homology with any known protein and no putative transmembrane domain was found. Interestingly, three different kinds of previously undescribed ORFs not homologous to each other were also found upstream to the ORF encoding RdRp in the segment 2 sequences of six otarine PBVs in this study (**Figure 2**). Similar to ORF1 and ORF2, these three ORFs were predicted by the gene prediction program FGENESV, instead of only the ORF Finder. These three kinds of ORFs showed no homology to any known proteins and did not possess any known protein domains, families or functional sites. Further experiments will be required to determine the functions of these three ORFs as well as ORF1 and ORF2.

Picobirnaviruse probably evolves through mechanisms similar to other segmented RNA viruses (**Figure 4**). The most thoroughly studied segmented RNA viruses are the influenza viruses (negative-sense single-stranded RNA virus) and rotaviruses (double-stranded RNA virus). Influenza viruses evolve through reassortment of RNA segments resulting in antigenic shifts and RNA mutations leading to antigenic drifts and causing major pandemics, lots of mortalities and morbidities, and economic lost (Smith et al., 2009). As for rotaviruses, virus strains belonging to the same group can also reassort their genomes resulting in enormous diversity (Kirkwood, 2010). This provides one of the mechanisms for the emergence of new rotavirus strains leading to disease outbreaks and loss of vaccine efficacy (Kirkwood, 2010). The presence of multiple segment 1 and segment 2 in the same animal observed in this study provides the PBV genome a good opportunity for reassortment (**Figure 4**). In addition, in the serially collected fecal samples from sea lions, it was also observed that two segment 1 sequences that belonged to the same clade (C3) can be present in the same animal (GiGi) from two consecutive years (**Table 4**). There were 18 nucleotide changes in these two segment 1 sequences, leading to 11 amino acid changes. This could be due to persistence of the same PBV with mutational changes over 2 years or re-infection by another otarine PBV with segment 1 of the same clade but of a different sequence type.

In this study, a high diversity of PBVs was observed in a variety of terrestrial and marine mammals. Multiple sequence types with significant differences, representing multiple strains of PBV, were present in the majority of PBV-positive samples from different kinds of animals. These results suggest that PBV probably evolves through mechanisms similar to other segmented RNA viruses.

AUTHOR CONTRIBUTIONS

PW conceived of the study, designed the study, contributed reagents and drafted the manuscript. JT conceived of the study, designed the study, participated in data analysis and drafted the manuscript; RB carried out the molecular lab work and participated in data analysis. AW and AT participated in data analysis; PM and S-WH contributed reagents; CCL, SA, CY, GC, KL, and CSL carried out the lab work. SL revised the manuscript and contributed reagents; K-YY conceived of the study, designed the study, contributed reagents and revised the manuscript. All authors gave final approval for publication.

FUNDING

This work is partly supported by the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Disease for the HKSAR Department of Health; Strategic Research Theme Fund, The University of Hong Kong; and Croucher Senior Medical Research Fellowship.

ACKNOWLEDGMENTS

We thank Christopher M. Riggs and members of Department of Veterinary Clinical Services, Hong Kong Jockey Club for assistance in collection of horse samples and Chung-Tong Shek, Agriculture, Fisheries and Conservation Department, HKSAR for assistance in collection of bat and monkey samples. We thank Tsz Ho Chiu for facilitation of the study.

REFERENCES

Attoui, H., Billoir, F., Cantaloube, J. F., Biagini, P., de Micco, P., and de Lamballerie, X. (2000). Strategies for the sequence determination of viral dsRNA genomes. *J. Virol. Methods* 89, 147–158. doi: 10.1016/S0166-0934(00)00212-3

Bányai, K., Martella, V., Bogdán, A., Forgách, P., Jakab, F., Meleg, E., et al. (2008). Genogroup I picobirnaviruses in pigs: evidence for genetic diversity and relatedness to human strains. *J. Gen. Virol.* 89, 534–539. doi: 10.1099/vir.0.83134-0

Banyai, K., Potgieter, C., Gellert, A., Ganesh, B., Tempesta, M., Lorusso, E., et al. (2014). Genome sequencing identifies genetic and antigenic divergence of porcine picobirnaviruses. *J. Gen. Virol.* 95, 2233–2239. doi: 10.1099/vir.0.057984-0

Bodewes, R., van der Giessen, J., Haagmans, B. L., Osterhaus, A. D., and Smits, S. L. (2013). Identification of multiple novel viruses, including a parvovirus and a hepevirus, in feces of red foxes. *J. Virol.* 87, 7758–7764. doi: 10.1128/JVI.00568-13

Chen, M., Sun, H., Lan, D., Hua, X., Cui, L., Yuan, C., et al. (2014). Molecular detection of genogroup I and II picobirnaviruses in pigs in China. *Virus Genes* 48, 553–556. doi: 10.1007/s11262-014-1058-8

Da Costa, B., Duquerroy, S., Tarus, B., and Delmas, B. (2011). Picobirnaviruses encode a protein with repeats of the ExxRxNxxxE motif. *Virus Res.* 158, 251–256. doi: 10.1016/j.virusres.2011.02.018

De Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., et al. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34, W362–W365. doi: 10.1093/nar/gkl124

Fregolente, M. C., de Castro-Dias, E., Martins, S. S., Spilki, F. R., Allegretti, S. M., and Gatti, M. S. (2009). Molecular characterization of picobirnaviruses from new hosts. *Virus Res.* 143, 134–136. doi: 10.1016/j.virusres.2009.03.006

Ganesh, B., Banyai, K., Masachessi, G., Mladenova, Z., Nagashima, S., Ghosh, S., et al. (2011a). Genogroup I picobirnavirus in diarrhoeic foals: can the horse serve as a natural reservoir for human infection? *Vet. Res.* 42:52. doi: 10.1186/1297-9716-42-52

Ganesh, B., Nagashima, S., Ghosh, S., Nataraju, S. M., Rajendran, K., Manna, B., et al. (2011b). Detection and molecular characterization of multiple strains of *Picobirnavirus* causing mixed infection in a diarrhoeic child: Emergence of prototype Genogroup II-like strain in Kolkata, India. *Int. J. Mol. Epidemiol. Genet.* 2, 61–72.

Ghosh, S., Kobayashi, N., Nagashima, S., and Naik, T. N. (2009). Molecular characterization of full-length genomic segment 2 of a bovine picobirnavirus (PBV) strain: evidence for high genetic diversity with genogroup I PBVs. *J. Gen. Virol.* 90, 2519–2524. doi: 10.1099/vir.0.013987-0

Gillman, L., Sanchez, A. M., and Arbiza, J. (2013). *Picobirnavirus* in captive animals from Uruguay: identification of new hosts. *Intervirology* 56, 46–49. doi: 10.1159/000338275

Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., and Pöhlmann, S. (2005). Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7988–7993. doi: 10.1073/pnas.0409465102

Kirkwood, C. D. (2010). Genetic and antigenic diversity of human rotaviruses: potential impact on vaccination programs. *J. Infect. Dis.* 202, S43–S48. doi: 10.1086/653548

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lau, S. K., Woo, P. C., Li, K. S., Huang, Y., Tsoi, H. W., Wong, B. H., et al. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14040–14045. doi: 10.1073/pnas.0506735102

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* 252, 1162–1164. doi: 10.1126/science.252.5009.1162

Malik, Y. S., Chandrashekar, K. M., Sharma, K., Haq, A. A., Vaid, N., Chakravarti, S., et al. (2011). *Picobirnavirus* detection in bovine and buffalo calves from foothills of Himalaya and Central India. *Trop. Anim. Health Prod.* 43, 1475–1478. doi: 10.1007/s11250-011-9834-0

Malik, Y. S., Kumar, N., Sharma, K., Sharma, A. K., Sircar, S., Jeena, L. M., et al. (2014). Molecular characterization of a genetically diverse bubaline *Picobirnavirus* strain, India. *Wetchasan Sattawaphaet* 39, 609–613.

Martinez, L. C., Masachessi, G., Carruyo, G., Ferreyra, L. J., Barril, P. A., Isa, M. B., et al. (2010). *Picobirnavirus* causes persistent infection in pigs. *Infect. Genet. Evol.* 10, 984–988. doi: 10.1016/j.meegid.2010.06.004

Nates, S. V., Gatti, M. S. V., and Ludert, J. E. (2011). The picobirnavirus: an integrated view on its biology, epidemiology and pathogenic potential. *Future Virol.* 6, 223–235. doi: 10.2217/fvl.10.76

Ng, T. F., Mesquita, J. R., Nascimento, M. S., Kondov, N. O., Wong, W., Reuter, G., et al. (2014). Feline fecal virome reveals novel and prevalent enteric viruses. *Vet. Microbiol.* 171, 102–111. doi: 10.1016/j.vetmic.2014.04.005

Pereira, H. G., Fialho, A. M., Flewett, T. H., Teixeira, J. M., and Andrade, Z. P. (1988a). Novel viruses in human faeces. *Lancet* 2, 103–104. doi: 10.1016/S0140-6736(88)90032-3

Pereira, H. G., Flewett, T. H., Candeias, J. A., and Barth, O. M. (1988b). A virus with a bisegmented double-stranded RNA genome in rat (*Oryzomys nigripes*) intestines. *J. Gen. Virol.* 69, 2749–2754. doi: 10.1099/0022-1317-69-11-2749

Ribeiro, S. R., Bezerra, D. A., Kaiano, J. H., Oliveira, D. S., Silvestre, R. V., Gabbay, Y. B., et al. (2014). Genogroup I avian picobirnavirus detected in Brazilian broiler chickens: a molecular epidemiology study. *J. Gen. Virol.* 95, 117–122. doi: 10.1099/vir.0.054783-0

Rosen, B. I., Fang, Z. Y., Glass, R. I., and Monroe, S. S. (2000). Cloning of human picobirnavirus genomic segments and development of an RT-PCR detection assay. *Virology* 277, 316–329. doi: 10.1006/viro.2000.0594

Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., et al. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459, 1122–1125. doi: 10.1038/nature08182

Smits, S. L., Schapendonk, C. M., van Beek, J., Vennema, H., Schurch, A. C., Schipper, D., et al. (2014). New viruses in idiopathic human diarrhea cases, the Netherlands. *Emerg. Infect. Dis.* 20, 1218–1222. doi: 10.3201/eid2007.140190

Symonds, E. M., Griffin, D. W., and Breitbart, M. (2009). Eukaryotic viruses in wastewater samples from the United States. *Appl. Environ. Microbiol.* 75, 1402–1409. doi: 10.1128/AEM.01899-08

Vabret, A., Mourez, T., Gouarin, S., Petitjean, J., and Freymuth, F. (2003). An outbreak of coronavirus OC43 respiratory infection in Normandy, France. *Clin. Infect. Dis.* 36, 985–989. doi: 10.1086/374222

van Leeuwen, M., Williams, M. M., Koraka, P., Simon, J. H., Smits, S. L., and Osterhaus, A. D. (2010). Human picobirnaviruses identified by molecular screening of diarrhea samples. *J. Clin. Microbiol.* 48, 1787–1794. doi: 10.1128/JCM.02452-09

Verma, H., Mor, S. K., Erber, J., and Goyal, S. M. (2015). Prevalence and complete genome characterization of turkey picobirnaviruses. *Infect. Genet. Evol.* 30, 134–139. doi: 10.1016/j.meegid.2014.12.014

Wakuda, M., Pongsuwanna, Y., and Taniguchi, K. (2005). Complete nucleotide sequences of two RNA segments of human picobirnavirus. *J. Virol. Methods* 126, 165–169. doi: 10.1016/j.jviromet.2005.02.010

Wang, Y., Banyai, K., Tu, X., and Jiang, B. (2012). Simian genogroup I picobirnaviruses: prevalence, genetic diversity, and zoonotic potential. *J. Clin. Microbiol.* 50, 2779–2782. doi: 10.1128/JCM.00634-12

Woo, P. C., Lau, S. K., Bai, R., Teng, J. L., Lee, P., Martelli, P., et al. (2012). Complete genome sequence of a novel picobirnavirus, otarine picobirnavirus, discovered in California sea lions. *J. Virol.* 86, 6377–6378. doi: 10.1128/JVI.00686-12

Woo, P. C., Lau, S. K., Chu, C. M., Chan, K. H., Tsoi, H. W., Huang, Y., et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* 79, 884–895. doi: 10.1128/JVI.79.2.884-895.2005

Woo, P. C., Lau, S. K., Lam, C. S., Tsang, A. K., Hui, S. W., Fan, R. Y., et al. (2014a). Discovery of a novel bottlenose dolphin coronavirus reveals a distinct species of marine mammal coronavirus in *Gammacoronavirus*. *J. Virol.* 88, 1318–1331. doi: 10.1128/JVI.02351-13

Woo, P. C., Lau, S. K., Teng, J. L., Tsang, A. K., Joseph, M., Wong, E. Y., et al. (2014b). Metagenomic analysis of viromes of dromedary camel fecal samples reveals large number and high diversity of circoviruses and picobirnaviruses. *Virology* 471, 117–125. doi: 10.1016/j.virol.2014.09.020

Yeager, C. L., Ashmun, R. A., Williams, R. K., Cardellichio, C. B., Shapiro, L. H., Look, A. T., et al. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* 357, 420–422. doi: 10.1038/357420a0

Zhang, S., Bai, R., Feng, R., Zhang, H., and Liu, L. (2015). Detection and evolutionary analysis of picobirnaviruses in treated wastewater. *Microb. Biotechnol.* 8, 474–482. doi: 10.1111/1751-7915.12239

Zhang, W., Li, L., Deng, X., Kapusinszky, B., Pesavento, P. A., and Delwart, E. (2014). The fecal virome of cats in an animal shelter. *J. Gen. Virol.* 95, 2553–2564. doi: 10.1099/vir.0.069674-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Woo, Teng, Bai, Wong, Martelli, Hui, Tsang, Lau, Ahmed, Yip, Choi, Li, Lam, Lau and Yuen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW
Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION
Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS
No limit to article distribution and re-use

TRANSPARENT
Editors and reviewers acknowledged by name on published articles

SUPPORT
By our Swiss-based editorial team

IMPACT METRICS
Advanced metrics track your article's impact

GLOBAL SPREAD
5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK
Our network increases readership for your article

Frontiers
EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on