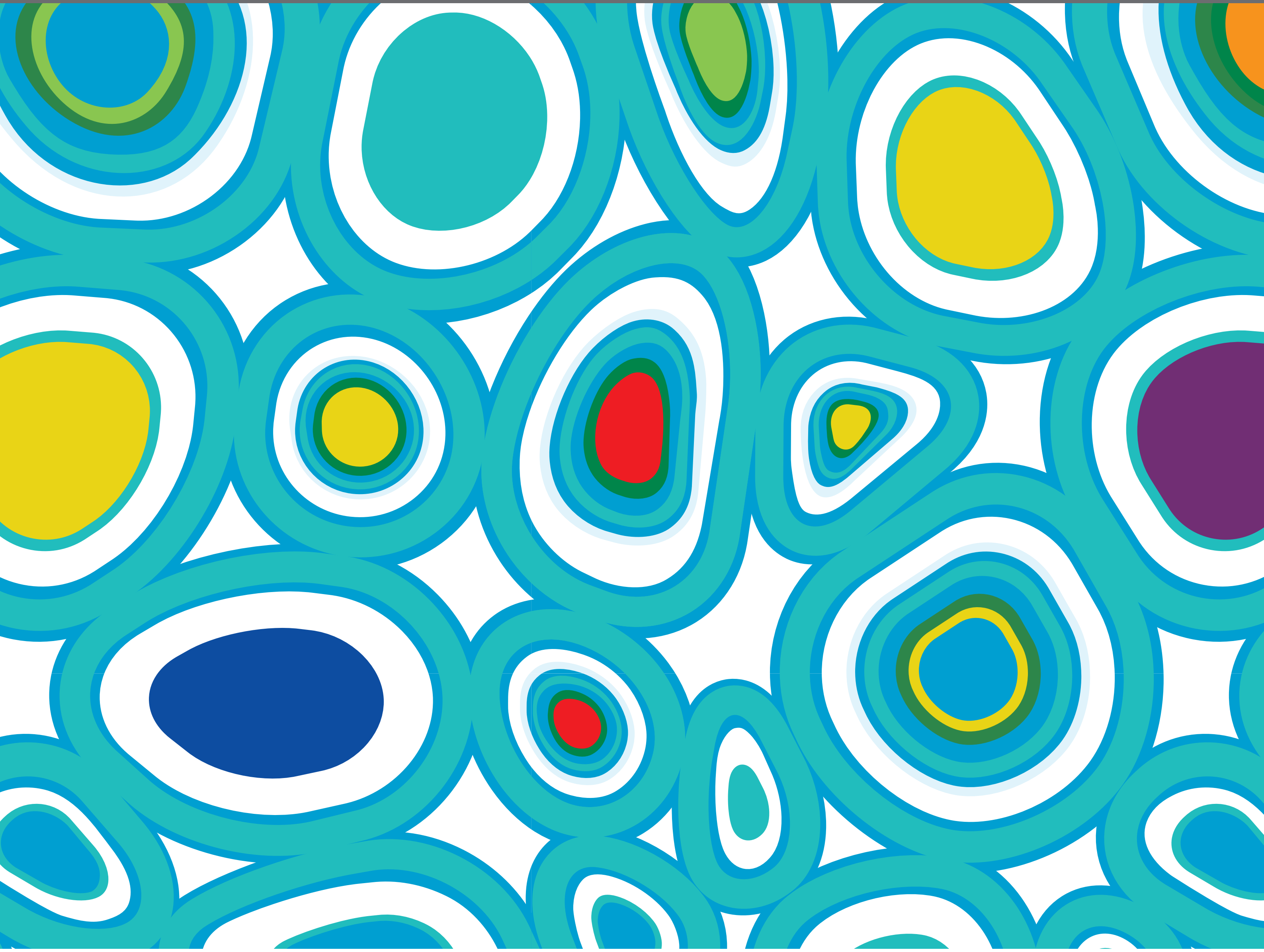
NEW APPROACHES IN CHORDATE AND VERTEBRATE EVOLUTION AND DEVELOPMENT

EDITED BY: Juan Pascual-Anaya, Stephanie Bertrand and
Salvatore D'Aniello
PUBLISHED IN: Frontiers in Cell and Developmental Biology and
Frontiers in Ecology and Evolution



frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88976-053-4
DOI 10.3389/978-2-88976-053-4

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: frontiersin.org/about/contact

NEW APPROACHES IN CHORDATE AND VERTEBRATE EVOLUTION AND DEVELOPMENT

Topic Editors:
Juan Pascual-Anaya, Malaga University, Spain
Stephanie Bertrand, UMR7232 Biologie Intégrative des Organismes Marins (BIOM), France
Salvatore D'Aniello, Zoological Station Anton Dohrn, Italy

Citation: Pascual-Anaya, J., Bertrand, S., D'Aniello, S., eds. (2022). New Approaches in Chordate and Vertebrate Evolution and Development. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-053-4

Table of Contents

05 ***Editorial: New Approaches in Chordate and Vertebrate Evolution and Development***
Juan Pascual-Anaya, Salvatore D'Aniello and Stephanie Bertrand

09 ***On the Embryonic Development of the Nasal Turbinals and Their Homology in Bats***
Kai Ito, Vuong Tan Tu, Thomas P. Eiting, Taro Nojiri and Daisuke Koyabu

28 ***Pescoids and Chimeras to Probe Early Evo-Devo in the Fish* Astyanax mexicanus**
Jorge Torres-Paz and Sylvie Rétaux

39 ***The Ontology of the Amphioxus Anatomy and Life Cycle (AMPHX)***
Stephanie Bertrand, João E. Carvalho, Delphine Dauga, Nicolas Matentzoglu, Vladimir Daric, Jr-Kai Yu, Michael Schubert and Hector Escrivá

48 ***An Updated Staging System for Cephalochordate Development: One Table Suits Them All***
João E. Carvalho, François Lahaye, Luok Wen Yong, Jenifer C. Croce, Hector Escrivá, Jr-Kai Yu and Michael Schubert

66 ***Prenatal Developmental Trajectories of Fluctuating Asymmetry in Bat Humeri***
Camilo López-Aguirre, Suzanne J. Hand, Daisuke Koyabu, Vuong Tan Tu and Laura A. B. Wilson

77 ***Innovation Through Heterochrony: An Amphioxus Perspective on Telencephalon Origin and Function***
Thurston Lacalli

81 ***Massive Gene Loss and Function Shuffling in Appendicularians Stretch the Boundaries of Chordate Wnt Family Evolution***
Josep Martí-Solans, Hector Godoy-Marín, Miriam Diaz-Gracia, Takeshi A. Onuma, Hiroki Nishida, Ricard Albalat and Cristian Cañestro

95 ***Evolution of Developmental Programs for the Midline Structures in Chordates: Insights From Gene Regulation in the Floor Plate and Hypochord Homologues of* Ciona *Embryos***
Kouhei Oonuma, Maho Yamamoto, Naho Moritsugu, Nanako Okawa, Megumi Mukai, Miku Sotani, Shuto Tsunemi, Haruka Sugimoto, Eri Nakagome, Yuichi Hasegawa, Kotaro Shimai, Takeo Horie and Takehiro G. Kusakabe

107 ***Improved Understanding of the Role of Gene and Genome Duplications in Chordate Evolution With New Genome and Transcriptome Sequences***
Madeleine E. Aase-Remedios and David E. K. Ferrier

122 ***Modular Evolution and Population Variability of* Oikopleura dioica *Metallothioneins***
Sara Calatayud, Mario Garcia-Risco, Mercè Capdevila, Cristian Cañestro, Òscar Palacios and Ricard Albalat

134 ***Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease***
Juan J. Tena and José M. Santos-Pereira

144 ***New Genes Born-In or Invading Vertebrate Genomes***
Carlos Herrera-Úbeda and Jordi Garcia-Fernàndez

153 ***A Preliminary Single-Cell RNA-Seq Analysis of Embryonic Cells That Express* Brachyury *in the Amphioxus,* Branchiostoma japonicum**
Noriyuki Satoh, Hitoshi Tominaga, Masato Kiyomoto, Kanako Hisata, Jun Inoue and Koki Nishitsuji

169 ***Optical Clearing and Light Sheet Microscopy Imaging of Amphioxus***
Simona Machacova, Helena Chmelova, Anna Vavrova, Zbynek Kozmik and Iryna Kozmikova

181 ***The Ascidia* Ciona robusta *Provides Novel Insights on the Evolution of the AP-1 Transcriptional Complex***
Pina Marotta, Federica Salatiello, Luca Ambrosino, Federica Berruto, Maria Luisa Chiusano and Annamaria Locascio

195 ***A Potential Method for Rapid Screening of Amphioxus Founder Harboring Germline Mutation and Transgene***
Jiaqi Zou, Xiaotong Wu, Chenggang Shi, Yanhong Zhong, Lei Zhang, Qiuning Yan, Liuru Su and Guang Li

204 ***Genetic Mechanism for the Cyclostome Cerebellar Neurons Reveals Early Evolution of the Vertebrate Cerebellum***
Fumiaki Sugahara, Juan Pascual-Anaya, Shigehiro Kuraku, Shigeru Kuratani and Yasunori Murakami

214 ***Brain Sensory Organs of the Ascidian* Ciona robusta*: Structure, Function and Developmental Mechanisms***
Paola Olivo, Antonio Palladino, Filomena Ristoratore and Antonietta Spagnuolo

226 ***Developmental Evolution of Hypaxial Muscles: Insights From Cyclostomes and Chondrichthyans***
Rie Kusakabe, Masako Tanaka and Shigeru Kuratani

235 ***When Bigger Is Better: 3D RNA Profiling of the Developing Head in the Catshark* Scyliorhinus canicula**
Hélène Mayeur, Maxence Lanoizelet, Aurélie Quillien, Arnaud Menuet, Léo Michel, Kyle John Martin, Sébastien Dejean, Patrick Blader, Sylvie Mazan and Ronan Lagadec

246 ***JNK Mediates Differentiation, Cell Polarity and Apoptosis During Amphioxus Development by Regulating Actin Cytoskeleton Dynamics and ERK Signalling***
Ildiko M. L. Somorjai, Matthias T. Ehebauer, Hector Escrivà and Jordi Garcia-Fernàndez

270 ***Derivedness Index for Estimating Degree of Phenotypic Evolution of Embryos: A Study of Comparative Transcriptomic Analyses of Chordates and Echinoderms***
Jason Cheok Kuan Leong, Yongxin Li, Masahiro Uesaka, Yui Uchida, Akihito Omori, Meng Hao, Wenting Wan, Yang Dong, Yandong Ren, Si Zhang, Tao Zeng, Fayou Wang, Luonan Chen, Gary Wessel, Brian T. Livingston, Cynthia Bradham, Wen Wang and Naoki Irie

288 ***Efficient CRISPR Mutagenesis in Sturgeon Demonstrates Its Utility in Large, Slow-Maturing Vertebrates***
Jan Stundl, Vladimír Soukup, Roman Franěk, Anna Pospisilova, Viktorie Psutkova, Martin Pšenička, Robert Cerny, Marianne E. Bronner, Daniel Meulemans Medeiros and David Jandzik

EDITORIAL
published: 12 May 2022
doi: 10.3389/fcell.2022.917101

Editorial: New Approaches in Chordate and Vertebrate Evolution and Development

Juan Pascual-Anaya[1,2]*, *Salvatore D'Aniello*[3] *and Stephanie Bertrand*[4]

[1]Department of Animal Biology, Faculty of Science, University of Málaga, Málaga, Spain, [2]Evolutionary Morphology Laboratory, RIKEN Cluster for Pioneering Research, Kobe, Japan, [3]Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Naples, Italy, [4]Sorbonne Université CNRS, Biologie Intégrative des Organismes Marins (BIOM), Observatoire Océanologique, Banyuls-sur-Mer, France

Keywords: chordate evolution, evo-devo, genomics, vertebrates, cephalochordates, urochordates (tunicates), transcriptomics, phenotypic evolution

Editorial on the Research Topic

New Approaches in Chordate and Vertebrate Evolution and Development

Consisting of three lineages—cephalochordates (amphioxus), urochordates (tunicates) and vertebrates (including us, humans) (**Figure 1**), the monophyletic group of chordate animals is defined by the presence, at some stage of their life cycle, of a set of unique and conserved morphological features: a dorsal hollow nerve cord with a notochord just ventral to it, pharyngeal slits and a post-anal tail. Moreover, the last common ancestor of chordates likely possessed a segmented muscular system along the main body axis (Satoh, 2016). Despite these synapomorphies, chordates have largely diversified since their origin around 600 million years ago, especially vertebrates, one of the most successful animal groups on our planet. The evolution of a myriad of morphological novelties (e.g., paired limbs, jaws) are behind chordate and vertebrate diversification since they permitted the adaptation of new species to a vast range of ecological niches.

Two processes are particularly relevant to chordate evolutionary success: the formation of the chordate body plan, and the origin of vertebrate traits. Traditionally, developmental biology and genetics have relied on the use of a handful of model animals (mostly jawed vertebrates like mouse, chicken, frog and fish and few invertebrate animals such as the ascidians *Ciona* and *Halocynthia*), offering a very narrow taxonomic range to tackle these two points. Although evolutionary developmental (evo-devo) studies soon attempted to increase the taxon sampling, these were impaired by the lack of accessible functional genomics techniques readily available for emerging model animals. Consequently, most of the early molecular genetics' studies in the field of chordate evo-devo consisted of the analysis of a single or a limited number of genes in few species. Therefore, and despite their importance, the mechanisms underlying the origin and evolution of both ancestral traits and morphological novelties of chordates have remained largely elusive. The recent surge and exponential growth of genome assemblies available from a wider set of animal models, together with the availability of the most recent technological innovations in genome editing and whole genome scale high throughput expression and regulation analyses are starting to provide answers to some of the long-standing chordate evo-devo questions.

With this in mind, we present this Frontiers Research Topic, entitled "*New Approaches in Chordate and Vertebrate Evolution and Development*," consisting of 23 manuscripts, including 14 original research articles and 5 reviews together with reports in several other formats. This fantastic Research Topic introduces multidisciplinary and integrative approaches that, by using the most recent state-of-the-art technologies in genomics, molecular biology, imaging and bioinformatics are contributing towards a better understanding of the mechanistic basis of chordate phenotypic diversity.

OPEN ACCESS

Edited by:
Ricard Albalat,
University of Barcelona, Spain

Reviewed by:
Pedro Martinez,
University of Barcelona, Spain

****Correspondence:***
Juan Pascual-Anaya
jpascualanaya@gmail.com

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *10 April 2022*
Accepted: *26 April 2022*
Published: *12 May 2022*

Citation:
Pascual-Anaya J, D'Aniello S and Bertrand S (2022) Editorial: New Approaches in Chordate and Vertebrate Evolution and Development.
Front. Cell Dev. Biol. 10:917101.
doi: 10.3389/fcell.2022.917101

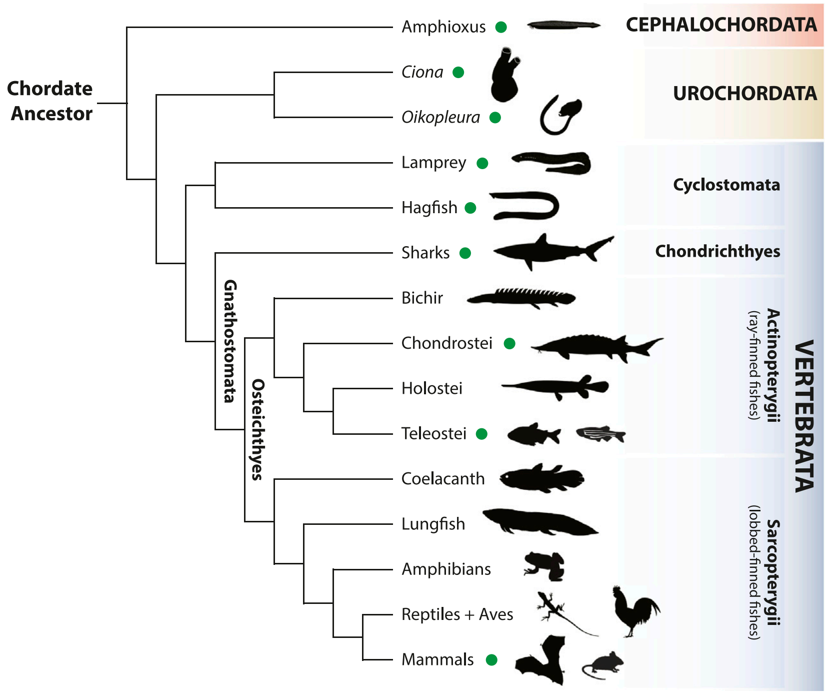


FIGURE 1 | Simplified phylogenetic tree of chordates (vertebrates, urochordates and cephalochordates). Models used in this Research Topic are marked with a green dot. Animal icons are from http://phylopic.org/ under a Creative Commons license.

Comparative genomics approaches are instrumental to decipher the genomic changes underpinning phenotypic evolution. While vertebrates have traditionally occupied the top position in terms of number and quality of genome sequencing projects, the recent burst of high-quality genome assemblies of less-studied chordate models have provided new opportunities to study genomic evolution at an unprecedented resolution. Aase-Remedios and Ferrier provide a thorough and updated view on how these new genomics and transcriptomics resources, especially in the vertebrate lineage, but also in key invertebrate chordates, have helped us to resolve long-standing questions, such as the exact phylogenetic timing of whole genome duplication events that took place during early vertebrate evolution, or the implications that gene duplicates have had in the evolution of developmental processes. In another review, Herrera-Úbeda and Garcia-Fernàndez address the fundamental question of what is the origin of new genes. They provide an interesting survey on how other processes beyond gene and genome duplication and divergence, such as horizontal gene transfer and the domestication of viruses and transposons, can supply substrates for the evolution of new genes and regulatory sequences, and how the integration of these into existing gene regulatory networks, when transcription factors or transcription factor binding sites are involved, are behind the origin of key vertebrate morphological novelties such as the placenta or the adaptive immune system. Leong et al. take advantage of transcriptomics resources available from different species of chordates, including its sister group, the echinoderms, to measure the degree of derivedness of developmental traits in evolutionary time. To address this question, they conceive a new method called "derivedness index", which is less constrained than other methods (which use exclusively conserved genomic traits) by taking into consideration lost or duplicated genes (and therefore non conserved), thus permitting to measure the degree of phenotypic evolution through more comprehensive comparative transcriptomics analyses. Tena and Santos-Pereira provide a review on how new technological advances on chromatin conformation studies have enabled us to better understand how the three-dimensional nature of genomes constrain the regulatory landscape of genes, and how changes in these 3D domains can lead to pathological situations, but at the same time provides opportunities for evolutionary innovation.

Cephalochordates, or amphioxus, share with tunicates and vertebrates the oldest common ancestor of chordates. This key phylogenetic position in the chordate tree (**Figure 1**), together with its slow evolving nature, both at the morphological and genomic levels, turn amphioxus into a unique and crucial animal for understanding the origin of chordate traits. Two original articles now provide new unvaluable resources for comparative developmental studies using amphioxus embryos. First, Carvalho et al., based on confocal imaging of embryological stages of the European amphioxus *Branchiostoma lanceolatum*, used as reference, propose a universal staging system for lancelet species, which they corroborate with comparisons with other amphioxus species. They also propose an unambiguous nomenclature of developmental stages, which will undoubtedly standardize the use of amphioxus embryos from different species across studies. A further extension of this universal staging system for lancelet development is put forward by Bertrand et al., who introduce the first ontology resource for an amphioxus species, termed AMPH, with ontology terms classified as developmental and anatomical specific entities, and hosted in open, public repositories. Importantly, these two amphioxus studies will further facilitate the use of amphioxus in large comparative analysis in evo-devo studies. Lacalli reviews very recent gene expression data on the most anterior region of the central nervous system of amphioxus and provides a thought-provoking article describing the role of heterochronic shifts in the origin of the telencephalon in the vertebrate brain, suggesting that the olfactory function is the driver of such shift. Somorjai et al. show, through careful pharmacological inhibition experiments, the roles that JNK (c-Jun N-terminal kinase) exerts during amphioxus development. Specifically, they find that JNK is particularly important in gastrulation, notochord formation and antero-posterior elongation of the embryo. Also on this Research Topic, Satoh et al. provide the first single-cell RNA-seq analysis in a cephalochordate, the Japanese lancelet *Branchiostoma japonicum*. They generated this preliminary resource from six different amphioxus stages and used it to better understand the expression divergence of the two Brachyury genes present in the amphioxus genome, *Bra1* and *Bra2*. Two last articles focusing on amphioxus provide new methodologies that promise to significantly advance our knowledge of cephalochordates and their use in chordate evo-devo studies. First, Machacova et al. describe a method that combines tissue clearing and light sheet microscopy, which significantly improves the rapid observation of whole-mount post-metamorphic amphioxus, from 1 to 6 month-old stages. Last, Zou et al. provide a simple method for the early detection of mutant and transgenic founders of amphioxus carrying either desired mutations–introduced by genome editing techniques such as the CRISPR/Cas9 technology—or transgenes–introduced by the Tol2 transposase system—, respectively, demonstrating their suitability to be used in amphioxus. This method promises to significantly increase the use of these state-of-the-art technologies for functional genomics studies in amphioxus, enabling a better understanding of the mechanistic processes of chordate development.

Urochordates, or tunicates, are the sister group of vertebrates (**Figure 1**) and, as such, can provide unvaluable insights into vertebrate origins. Our Research Topic contains two outstanding studies focusing on the fast-evolving appendicularian *Oikopleura dioica*, which represents a unique case to study evolutionary divergence in chordates due to its reduced genome size and large amount of gene losses. The first of these studies, from Calatayud et al., investigates metallothionein proteins, a family of modular proteins with a variable number of cysteine-rich domains. The different number of domains in different genes entail these proteins different binding capacities to transition metal ions. The authors also found that variants of these proteins with different numbers of repetitive domains exist in distinct populations of *O. dioica*. Their work offers a new model in which to study protein modular evolution. Then, Martí-Solans et al. highlight the importance of gene losses and their implications in evolution. These last authors analyzed the Wnt (wingless) gene family, their numbers and expression patterns, finding that few Wnt genes have been retained in *O. dioica*, which therefore contains the shortest, yet functional, Wnt repertoire described so far in chordates. Other three studies focus on the more traditional tunicate model, the ascidian *Ciona* spp.: Marotta et al. identified the gene members of the Activator Protein-1 transcription factor family (AP-1) and their expression patterns in *Ciona robusta*. Given that only single copies of these genes exist in *Ciona*, this enables *Ciona* as a helpful experimental organism to study the evolution of the AP-1 transcriptional complex in chordates; Oonuma et al. provide evidence of a common regulatory network involving *hedgehog* and *FoxA* genes in the origin of midline structures in tunicates and vertebrates, namely the floor plate of the neural tube, plus the hypochord in vertebrates and the endodermal strand in tunicates; last, Olivo et al. review the molecular bases of the pigment cells and the otolith of *C. robusta* in order to understand the evolution of sensory organs in chordates.

Vertebrates represent the most diverse subphylum of chordates (**Figure 1**), both in terms of morphological divergence as well as genome complexity. The vertebrate lineage is also characterized by key innovations such as the neural crest cells, a craniate head, and the jaw and paired appendages in the gnathostome lineage (Shimeld and Holland, 2000; Holland et al., 2015). Several articles in this Research Topic provide new insights into the evolution of vertebrates, focusing on all major vertebrate branches. In the cyclostome lineage (extant agnathans, or jawless vertebrates), Sugahara et al. report a comparative analysis of gene expression patterns in the dorsal side of the rhombomere 1 in both the lamprey and in the less accessible hagfish. They provide an updated evolutionary scenario for the origin of the cerebellum, involving the presence of a *Ptf1a-Atoh1* axis for the specification of inhibitory and excitatory neurons. Kusakabe et al. provide a review of the most recent data in lampreys and sharks concerning the evolution of a type of skeletal muscle, the hypaxial muscles, which are involved in the origin of key novel structures in vertebrates such as the fins and limbs. They propose a new evolutionary hypothesis in which the divergent expression of *Lbx* paralogs is behind the gradual elaboration of skeletal musculature during evolution. Then, Mayeur et al. present an innovative 3D atlas of expression data in the embryonic head of the catshark *Scyliorhinus canicula*. This important adaptation of RNA tomography in a non-model system will bring useful information to other researchers that aim to apply this technique to different chordate species. Continuing with the adaptation of state-of-the-art technologies to emerging experimental organisms, Stundl et al. describe a method for the use of the CRISPR/Cas9 system into a non-model, slow-growing, non-teleost fish: the sterlet sturgeon *Acipenser ruthenus*. Sturgeons belong to the chondrosteans, an important lineage in evolutionary studies of fishes due to its key phylogenetic position as the sister group to two major lineages of ray-finned fishes: holosteans and teleosts. In a proof-of-concept assay, the authors applied the method to successfully mutate the *Tyrosinase* and *Sonic hedgehog* genes. In another research article, Torres-Paz and Rétaux present a powerful tool for experimental embryology, describing a method to grow blastoderm explants (pescoids) and to produce chimeric embryos of the Mexican tetra, *Astyanax mexicanus*, an important evo-devo experimental organism to understand phenotypic evolution. Using these methods, they studied the impact of embryonic and extra-embryonic tissues on cell fate decisions and early developmental processes. To finish, this Research Topic also includes two original articles on the study of bat evolution: first, Ito et al. use advanced computed tomography, called diceCT (from diffusible iodine-based contrast-enhanced computed tomography), to study in detail the 3D prenatal development of various bat species, providing new insights into the evolution of nasal turbinals in bats; and lastly, López-Aguirre et al. studied the development of the humerus in bats, a very specialized bone, using a geometric morphometrics approach to find the presence of fluctuating asymmetry, both on the longitudinal and cross-sectional patterns of humeri, which might be under developmental control of independent mechanisms. They discuss the implications their results might have to understand the selective pressures behind this structure, which is crucial for the specific ecological adaptations of bats.

In conclusion, the collection of articles in this Research Topic introduces the most-updated current state of evo-devo research in chordates, presenting original studies in a wide taxonomic range of animals, and also providing novel methodologies that we hope will significantly improve studies in other evo-devo experimental systems.

AUTHOR CONTRIBUTIONS

JP-A was the Guest editor of this Research Topic, inviting co-editors SD'A and SB. All authors acted as handling editors of manuscripts in the Research Topic. JP-A wrote the Editorial with input from the other co-editors. All authors approved the submitted version.

FUNDING

JP-A holds a Beatriz Galindo position (BEAGAL20/00118) from the "Ministerio de Ciencia, Innovación y Universidades" of Spain.

ACKNOWLEDGMENTS

We would like to thank all the authors who have contributed to this Research Topic. We also feel indebted to all the reviewers for their essential contribution to this Research Topic.

REFERENCES

Holland, N. D., Holland, L. Z., and Holland, P. W. H. (2015). Scenarios for the Making of Vertebrates. *Nature* 520 (7548), 450–455. doi:10.1038/nature14433

Satoh, N. (2016). "Deuterostomes and Chordates," in *Chordate Origins and Evolution*. Editor N. Satoh (San Diego: Academic Press), 1–16. doi:10.1016/b978-0-12-802996-1.00001-8

Shimeld, S. M., and Holland, P. W. H. (2000). Vertebrate Innovations. *Proc. Natl. Acad. Sci. U.S.A.* 97 (9), 4449–4452. doi:10.1073/pnas.97.9.4449

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Pascual-Anaya, D'Aniello and Bertrand. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 23 March 2021
doi: 10.3389/fcell.2021.613545

On the Embryonic Development of the Nasal Turbinals and Their Homology in Bats

Kai Ito[1]*, *Vuong Tan Tu*[2,3], *Thomas P. Eiting*[4], *Taro Nojiri*[5,6] *and Daisuke Koyabu*[7,8,9]*

[1] *Department of Anatomy, Tissue and Cell Biology, School of Dental Medicine, Tsurumi University, Yokohama, Japan,* [2] *Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Hanoi, Vietnam,* [3] *Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam,* [4] *Department of Neurobiology and Anatomy, University of Utah, Salt Lake City, UT, United States,* [5] *Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan,* [6] *The University Museum, The University of Tokyo, Tokyo, Japan,* [7] *Research and Development Center for Precision Medicine, University of Tsukuba, Tsukuba, Japan,* [8] *Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Kowloon, Hong Kong,* [9] *Department of Molecular Craniofacial Embryology, Tokyo Medical and Dental University, Tokyo, Japan*

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
RIKEN Cluster for Pioneering
Research (CPR), Japan

Reviewed by:
Oleksandr Yaryhin,
Max Planck Institute for Evolutionary
Biology, Germany
Sérgio Ferreira-Cardoso,
UMR 5554 Institut des Sciences
de l'Evolution de Montpellier (ISEM),
France

***Correspondence:**
Daisuke Koyabu
dsk8evoluxion@gmail.com
Kai Ito
ocean42.rhino@gmail.com

Specialty section:
This article was submitted to
Evolutionary Developmental Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: *02 October 2020*
Accepted: *08 February 2021*
Published: *23 March 2021*

Citation:
Ito K, Tu VT, Eiting TP, Nojiri T and
Koyabu D (2021) On the Embryonic
Development of the Nasal Turbinals
and Their Homology in Bats.
Front. Cell Dev. Biol. 9:613545.
doi: 10.3389/fcell.2021.613545

Multiple corrugated cartilaginous structures are formed within the mammalian nasal capsule, eventually developing into turbinals. Due to its complex and derived morphology, the homologies of the bat nasal turbinals have been highly disputed and uncertain. Tracing prenatal development has been proven to provide a means to resolve homological problems. To elucidate bat turbinate homology, we conducted the most comprehensive study to date on prenatal development of the nasal capsule. Using diffusible iodine-based contrast-enhanced computed tomography (diceCT), we studied in detail the 3D prenatal development of various bat species and non-bat laurasiatherians. We found that the structure previously identified as "maxilloturbinal" is not the true maxilloturbinal and is only part of the ethmoturbinal I pars anterior. Our results also allowed us to trace the evolutionary history of the nasal turbinals in bats. The turbinate structures are overall comparable between laurasiatherians and pteropodids, suggesting that pteropodids retain the ancestral laurasiatherian condition. The absence of the ethmoturbinal I pars posterior in yangochiropterans and rhinolophoids has possibly occurred independently by convergent evolution.

Keywords: Chiroptera, evo-devo, skull, microCT (μCT), homology

INTRODUCTION

The mammalian nasal cavity contains a series of bony and cartilaginous plate-like structures called turbinals, which together project into the nasal cavity and provide surface area for various functions (Moore, 1981; Van Valkenburgh et al., 2014; Smith et al., 2015). Generally, the roles of the nasal cavity are twofold: to heat and humidify inhaled air before entering the lungs and to provide surface area for odorant deposition and olfactory sensation (Negus, 1958; Hillenius, 1992). The turbinals projecting into the nasal cavity primarily provide a surface area, offering a scaffold for blood vessels, secretory cells, and olfactory cells (Negus, 1958; Moore, 1981; Smith and Rossie, 2008; Van Valkenburgh et al., 2011; Wagner and Ruf, 2019, 2020).

Several types of turbinal can be recognized in the mammalian nasal cavity. The marginoturbinal and atrioturbinal are found in the outer nasal cartilage in the rostral part of the nasal cavity. The marginoturbinal begins at the lateral margin of the external nasal opening and continues into the atrioturbinal (Maier, 1980, 2000). The shape of these turbinals forms the naris and permits effective airflow (Göbbel, 2000; Maier and Ruf, 2014). These turbinals remain cartilaginous in adults and are continuous with the maxilloturbinal (Voit, 1909; Reinbach, 1952a,b; Maier, 1980, 1993b; Zeller, 1987; Smith et al., 2015). The maxilloturbinal is ventrally positioned in the nasal cavity (Negus, 1958; Moore, 1981; Smith et al., 2015). This turbinal projects from the medial surface of the maxilla and is covered with the respiratory epithelium (Scott, 1954; Adams, 1972; Van Valkenburgh et al., 2011) to add humidity to and increase the temperature of inhaled air (Scott, 1954). The maxilloturbinal generally becomes the largest and most complex in adults (Maier, 1993b; Maier and Ruf, 2014; Van Valkenburgh et al., 2014; Smith et al., 2015). The nasoturbinal projects from the roof of the nasal cavity (Moore, 1981). This turbinal articulates with the inferior margin of the nasal bone and medial surface of the maxilla and extends caudally into the ethmoid complex (Moore, 1981; Smith et al., 2015). The ethmoturbinals project from the lateral mass of the ethmoid bone (Smith et al., 2015). Several ethmoturbinals are found (Van Gilse, 1927; Maier, 1993a; Maier and Ruf, 2014) and are generally covered with olfactory epithelium (Adams, 1972; Gross et al., 1982; Martinez et al., 2020). Each ethmoturbinal is arranged one behind the other in parallel (Voit, 1909). Voit (1909) denoted the ethmoturbinals by Roman numerals in rostrocaudal sequence. Ethmoturbinal I protrudes toward the nostrils and is usually the largest among the ethmoturbinals (Voit, 1909). It makes the front border of the ethmoturbinal recess, which is the restricted space in the caudal part within the nasal cavity (Smith and Rossie, 2008; Maier and Ruf, 2014). The number of ethmoturbinals varies among species (Paulli, 1900a,b,c; Rowe et al., 2005). To our knowledge, the minimum number is seen in *Tursiops* in odontocetes with the absence of the ethmoturbinal (Mead and Fordyce, 2009). *Orycteropus afer* has the maximum number of ethmoturbinal so far with "at least nine" (Stößel et al., 2010). The frontoturbinals are located within the frontoturbinal recess, which is the dorsocaudal space of the lateral recess bounded ventrally by the root of ethmoturbinal I (Maier, 1993a; Rossie, 2006). The accessory scrolls between the frontoturbinals within the frontoturbinal recess are known as interturbinals (Maier, 1993b; Maier and Ruf, 2014; Ruf, 2014). The number of frontoturbinals and interturbinals may vary depending on the species (Smith et al., 2015).

In addition to turbinals, the nasal cavity presents other sheet-like ossifications such as the lamina semicircularis, lamina horizontalis, and lamina transversalis (Maier and Ruf, 2014). The lamina semicircularis is the medial wall of the maxillary recess and frontoturbinal recess (Ruf, 2014). This lamina is continuous with the posterior part of the nasoturbinal (Macrini, 2012; Smith et al., 2015). The lamina horizontalis separates the lateral recess into the dorsal and ventral chambers: the dorsal chamber is the frontoturbinal recess and the ventral chamber is the maxillary recess (Smith and Rossie, 2008; Maier and Ruf, 2014). The lamina transversalis extends from the lateral walls of the nasal cavity and attaches to the nasal septum, separating the ethmoturbinal recess from the nasopharyngeal duct (Lozanoff and Diewert, 1989; Macrini, 2012; Smith et al., 2015).

As for the general developmental pattern for mammals, initially, these turbinals and laminae appear as simple ridges along the lateral wall of the nasal capsule (Dieulafe, 1906). The nasal capsule, which is the rostral part of the chondrocranium, undergoes drastic morphological changes through ontogeny (Maier and Ruf, 2014; Van Valkenburgh et al., 2014; Smith et al., 2015). Morphogenesis of the nasal capsule in mammals is attributed to three mesenchymal condensations: the parietotectal cartilage aside from the tectum, paranasal cartilage, and orbitonasal lamina (De Beer, 1937; Reinbach, 1952b; Moore, 1981; Zeller, 1987; Rossie, 2006; Smith and Rossie, 2006, 2008; Van Valkenburgh et al., 2014). The nasal tectum of the parietotectal cartilage condenses in the rostrocaudal and mediolateral direction (De Beer, 1937; Smith and Rossie, 2008). As the mesenchyme condenses, the rostral ridge of the paranasal cartilages overlaps the parietotectal cartilage, and the caudal ridge of the paranasal cartilages overlaps the orbitonasal lamina (Smith and Rossie, 2008; Van Valkenburgh et al., 2014). As a result, the lamina semicircularis is formed rostrally and ethmoturbinal I is formed caudally within the nasal capsule. Subsequently, ethmoturbinals II to IV are formed rostrocaudally within the orbitonasal lamina (Rossie, 2006; Smith and Rossie, 2008; Van Valkenburgh et al., 2014). The nasal capsule then becomes gradually enclosed by exocranial facial bones (Maier and Ruf, 2014). Through prenatal ontogeny, the structure of each turbinal changes in shape and becomes complicated, filling the nasal cavity (Maier and Ruf, 2014). Prenatally, the nasal epithelium sinks at specific sites, where the initial folds are created. Within the initial folds, mesenchymal condensations constitute the primitive morphology of the chondral template of turbinals (Smith et al., 2020). Later, these mesenchymal condensations chondrify. In perinatal and postnatal stages, cartilages change their shape into lamellae (Smith et al., 2020). In the adult, the cartilaginous lamellae is fully ossified with the process of endochondral ossification except for the marginoturbinal and atrioturbinal (Voit, 1909; Martineau-Doizé et al., 1992; Ruf et al., 2015; Smith et al., 2020). Additional turbinals branch off from each turbinal, scroll, and fold and also merge with one another (Parker, 1885; Maier, 1980, 2000; Deleon and Smith, 2014; Maier and Ruf, 2014; Smith et al., 2020). The ossified remnant of the nasal capsule becomes the ethmoid bone (Patterson, 1977). An emerging consensus agrees with the bauplan (body plan) of cartilaginous nasal capsule having a tripartite composition: the anterior part (pars anterior), lateral part (pars lateralis), and posterior part (pars posterior) (**Figure 1**) (Maier, 1993a; Rossie, 2006; Smith and Rossie, 2006, 2008; Maier and Ruf, 2014; Van Valkenburgh et al., 2014).

The turbinate anatomy of various mammalian species at the adult stage has been described by the classic works of Paulli, providing a major source of current information on the diversity of mammalian turbinals (Paulli, 1900a,b,c). However, his studies erroneously interpreted the lamina semicircularis as a turbinal, due to the lack of observations on fetal stages of nasal structures

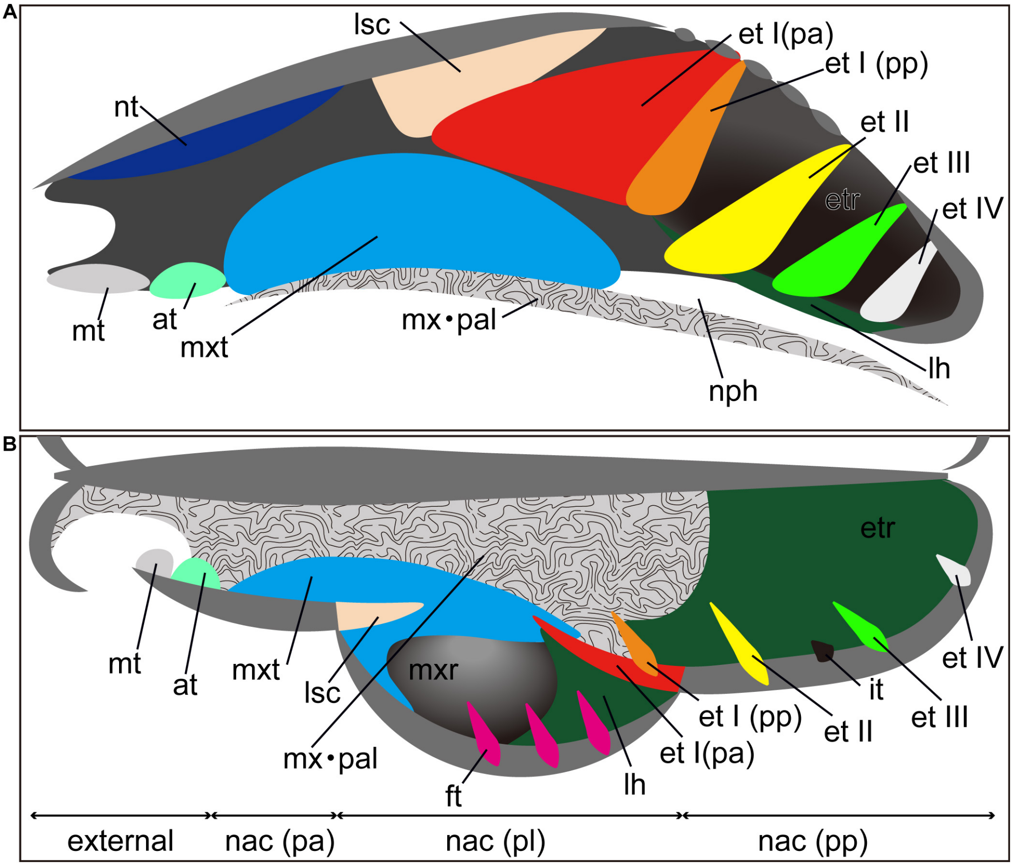


FIGURE 1 | Generalized schematic mammalian nasal capsule (modified from Maier, 1993a). **(A)** Medial view of parasagittal section; **(B)** horizontal section. These images show the nasal structure without facial exocranial (dermal) bones except for the maxilla and palatine bones. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pp), ethmoturbinal I pars posterior; et II–IV, ethmoturbinals II–IV; etr, ethmoturbinal recess; ft, frontoturbinal; in, interturbinal; lh, lamina horizontalis; lsc, lamina semicircularis; mt, marginoturbinal; mx, maxillare; mxr, maxillary recess; mxt, maxilloturbinal; nac (pa), nasal capsule pars anterior; nac (pl), nasal capsule pars lateralis; nac I (pp), nasal capsule pars posterior; nph, nasopharyngeal duct; nt, nasoturbinal; pal, palatinum.

(Maier and Ruf, 2014). Since the nasal structure becomes highly complicated, particularly during prenatal development, it is virtually impossible to correctly establish turbinate homologies between species solely by comparisons of adult anatomy (Maier and Ruf, 2014). In contrast, the turbinate structure in fetal stages is rather simple, and observations on fetal series allow us for tracing the structural changes of the nasal capsule and turbinals (Maier, 1993a; Macrini, 2014; Maier and Ruf, 2014). Thus, previous studies have emphasized the importance of comparative embryological approaches for understanding turbinate homology among mammals (Novacek, 1993; Maier, 1993a; Maier and Ruf, 2014). However, few studies incorporate fetal samples (Reinbach, 1952a,b; Maier, 1980; Smith and Rossie, 2008; Giannini et al., 2012; Maier and Ruf, 2014; Ruf et al., 2015; Ruf, 2020), possibly due to the difficulty in obtaining rare fetal samples.

Bats lack the prenatal information on turbinate anatomy with unresolved turbinate homology. They are the second most speciose order of mammals, exceeding 1,400 recognized species (Wilson and Mittermeier, 2019; Simmons and Cirranello, 2020). Phylogenetically, they are presently divided into two suborders, i.e., Yangochiroptera and Yinpterochiroptera (Springer et al., 2001; Teeling et al., 2002, 2003, 2005). Apart from most members of the family Pteropodidae of the Yinpterochiroptera, many bat species can use laryngeal echolocation. Most echolocating bat species emit their calls orally, but in some families, echolocation calls are emitted nasally (Brigham et al., 2004; Feldhamer et al., 2007). Olfactory capabilities in bats have been suggested to vary between species (Bhatnagar and Kallen, 1974a; Hutcheon et al., 2002). Bat turbinals have been studied by many authors (Grosser, 1900; Frick, 1954; Bhatnagar and Kallen, 1974a,b, 1975; Cooper and Bhatnagar, 1976; Göbbel, 2002; Giannini et al., 2006, 2012; Nelson et al., 2007; Smith et al., 2012; Eiting et al., 2014a; Curtis and Simmons, 2017; Curtis et al., 2020; Yohe et al., 2018), but the complex and diverse anatomy of bat turbinals has caused much confusion regarding their homology, possibly owing to the variations in echolocation behavior and olfactory functions (Curtis and Simmons, 2017; Curtis et al., 2020).

A handful of studies have attempted to discuss the homologies of bat turbinals (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Curtis and Simmons, 2017) using adult specimens; however, as noted earlier, homologies of the

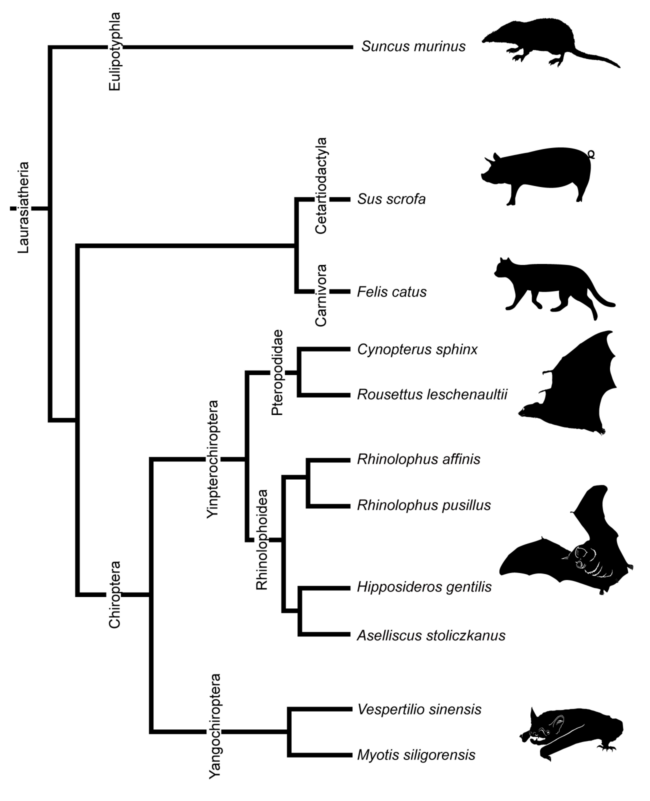


FIGURE 2 | Phylogenetic relationships of bats and outgroup species in this study. Phylogenetic framework is based on Teeling et al. (2005) and Li et al. (2007).

mammalian nasal structures are hardly possible to establish without studying fetal anatomy. To date, our knowledge on prenatal turbinals in bats is still in its infancy and restricted to only a few studies on some bat species, including *Rousettus aegyptiacus* (Jurgens, 1962; Fehse, 1990), *Pipistrellus pipistrellus*, *Rhinolophus ferrumequinum*, *Vespertilio murinus* (Grosser, 1900), *Miniopterus schreibersii* (Fawcett, 1919; De Beer, 1937), *Myotis myotis* (Frick, 1954), *Pteropus lylei* (Giannini et al., 2012), *Megaderma lyra* (Smith et al., 2012), and *Rousettus leschenaultii* (Smith et al., 2020).

Similarly, the fetal anatomy and ontogenetic periods to adult stages in bats are still largely unexplored or poorly studied. Here, using diffusible iodine-based contrast-enhanced computed tomography (diceCT) imaging, we describe the detailed embryonic development of the nasal cavity in eight species of bats, dividing into two suborders: Yangochiroptera and Yinpterochiroptera. We revise turbinate homologies among bats and reconstruct the evolutionary history of the nasal turbinal of bats in light of the modern phylogenetic framework.

MATERIALS AND METHODS

We observed multiple developmental stages from the fetus to adult of eight species of bats. Stages and basic measurements are summarized in **Supplementary Table 1**. Our samples include *Cynopterus sphinx* and *R. leschenaultii* from Pteropodidae, *Rhinolophus affinis* and *Rhinolophus pusillus* from Rhinolophidae, *Hipposideros gentilis* and *Aselliscus stoliczkanus* from Hipposideridae of Yinpterochiroptera, and *Myotis siligorensis* and *Vespertilio sinensis* from Vespertilionidae species of Yangochiroptera (**Figure 2**). Fetuses of three non-bat species of Laurasiatheria, *Suncus murinus*, *Sus scrofa*, and *Felis catus*, were included as outgroups (**Figure 2**). Samples belong to the curatorial collections at the Institute of Ecology and Biological Resources of Vietnam Academy of Science and Technology and the University Museum of the University of Tokyo (**Supplementary Table 1**). These samples were fixed and preserved with 70% ethanol solution. Grayscale images of the specimens' crania were obtained using microCT (InspeXio SMX-90CT Plus, Shimadzu Co, Japan) with 90 kV source voltage and 100 mA source currents. To enhance the contrast of the CT images, we followed the image enhancement techniques of a previous study (Gignac and Kley, 2014; Gignac et al., 2016) and dipped the specimens with iodine-based solutions (1% iodine, I_2KI in 99% ethanol solution) (Sohn et al., in press). Staining duration was between 6 and 24 h depending on the size of the specimen. Voxel size ranged from 8 to 35 μm. Images were reconstructed with dimensions of 1,024 × 1,024 pixels and in 12-bit grayscale. We reconstructed the cartilage and bones within turbinals by manual segmentation of grayscale images for each specimen using Segmentation Editor Tool in Amira 5.3 (Visage Imaging, Berlin, Germany) (**Supplementary Table 1**). The cartilaginous structures are stained poorly by iodine-based solutions. We identified them indirectly from the connective tissue like perichondria, which are readily stained with iodine-based solutions (Gignac et al., 2016). When interpreted from the surrounding structure, it is possible to distinguish ossified and cartilaginous structures. **Supplementary Figure 1** shows the ossified and cartilaginous structure with enhanced contrast of the CT images from iodine solution. The crown-rump length (CRL) of each specimen was measured using sliding calipers (N20, Mitutoyo, Japan). Bat specimens were staged following Cretekos et al. (2005), which has been developed based on the Carnegie system for human development. Bat fetal specimens of stages CS18, CS19, and CS22 or CS23 of Cretekos' staging system (which respectively correspond to CS18, CS19, and CS22 in the human Carnegie system), in which turbinate development and splitting can be observed, were here compared. In this study, specimens assigned as stage 18 are hereafter referred to as "early stage," stage 19 as "mid stage," and stages 22 and 23 as "late stage" for simplification. For *R. pusillus*, a fetal specimen of stage 15 was additionally studied to observe the initial onset of the turbinate projection. Gestation day 29 and postnatal day 1 of *S. murinus* are respectively referred as "mid stage" and "late stage" (which roughly correspond to CS22 and CS23 in the human Carnegie system). Gestation day 28 and gestation day 40 of *S. scrofa* are respectively referred as "mid stage" and "late stage" (which roughly correspond to CS22 and CS23 in the human Carnegie system). Gestation day 38 and gestation day 49 of *F. catus* are referred as "mid stage" and "late stage," respectively (which roughly correspond to CS22 and CS23 in the human Carnegie system). *S. scrofa* and *F. catus* were aged

TABLE 1 | Terminology for turbinals and laminas.

Structure name	Synonyms from other authors
Marginoturbinal	–
Atrioturbinal	–
Maxilloturbinal	Inferior concha (Moore, 1981, p. 255)
Nasoturbinal	Nasoturbinal, mucosal part (Smith and Rossie, 2008), rostral nasoturbinal (Giannini et al., 2012)
Lamina semicircularis	Crista semicircularis (Voit, 1909), endoturbinal I (Paulli, 1900a,b,c; Moore, 1981), semicircular crest (Smith and Rossie, 2008), caudal nasoturbinal (Giannini et al., 2012), nasoturbinal osseous part (Smith et al., 2015)
Lamina horizontalis	Anterior root of ethmoturbinal I (De Beer, 1937), lateral root of ethmoturbinal I (Rossie, 2006), frontomaxillary septum (Smith and Rossie, 2008), lamina transversalis posterior (Macrini, 2014)
Ethmoturbinal I pars anterior	Endoturbinal I (Allen, 1882; Giannini et al., 2012), endoturbinal II (Paulli, 1900a,b,c; Moore, 1981), middle concha (Moore, 1981, p. 255), ethmoturbinals I (Smith and Rossie, 2008), endoturbinal I in adult (Macrini, 2014)
Ethmoturbinal I pars posterior	Ethmoturbinal I lobule (Allen, 1882), endoturbinal II, lower lamella (Paulli, 1900a,b,c; Moore, 1981), middle concha (Moore, 1981, p. 255), ethmoturbinals II (Smith and Rossie, 2008), endoturbinal I in adult (Macrini, 2014)
Ethmoturbinal II	Endoturbinal II (Allen, 1882; Giannini et al., 2012), endoturbinal III (Allen, 1882; Paulli, 1900a,b,c; Moore, 1981), superior concha (Moore, 1981, p. 255), ethmoturbinals III (Smith and Rossie, 2008), endoturbinal II in adult (Macrini, 2014)
Ethmoturbinal III	Endoturbinal III (Allen, 1882; Giannini et al., 2012), endoturbinal IV (Paulli, 1900a,b,c; Moore, 1981), highest concha (Moore, 1981, p. 255), ethmoturbinal IV (Smith and Rossie, 2008)
Interturbinal	Ectoturbinal (Allen, 1882; Paulli, 1900a,b,c; Moore, 1981; Giannini et al., 2012)
Frontoturbinal	Ectoturbinal (Allen, 1882; Paulli, 1900a,b,c; Moore, 1981; Giannini et al., 2012), ectoturbinal in adult (Macrini, 2014)

based on Evans and Sack (1973). Specimen ID, CRL, stages, and scanning parameters of all specimens are summarized in **Supplementary Table 1**.

Terminology

The anatomical terminology for turbinals varies between studies (**Table 1**), but here we adopted the bauplan proposed by Maier (1993a) and followed the anatomical terminology of Voit (1909) (**Figure 1**). We chose this terminology because it takes into account the topography, ontogeny, and homology of turbinal bones (Maier and Ruf, 2014).

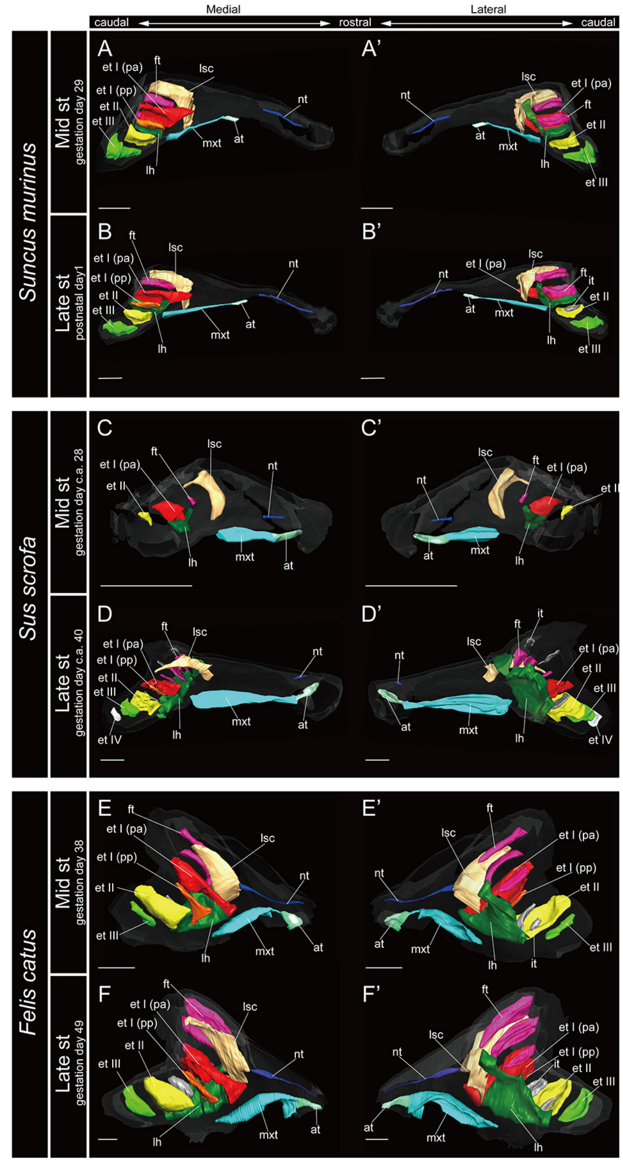


FIGURE 3 | 3D reconstructions of the developing turbinals in non-bat laurasiatherians. **(A,A′,C,C′,E,E′)** Mid stage fetus; **(B,B′,D,D′,F,F′)** late stage fetus or postnatal; **(A–F)** medial view; **(A′–F′)** lateral view. Scale bars, 1 mm. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pp), ethmoturbinal I pars posterior; et II–IV, ethmoturbinals II–IV; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; nt, nasoturbinal; lsc, lamina semicircularis.

RESULTS

Marginoturbinal and Atrioturbinal

The marginoturbinal and the atrioturbinal were cartilaginous structures in all species examined here. The atrioturbinal of all outgroup species and all bats was positioned ventrally, and it was continuous with the maxilloturbinal caudally (**Figures 3–7**). As the nasal capsule enlarged, the atrioturbinal became more rostrocaudally elongated in the outgroup species as well as in Pteropodidae (**Figures 3, 4**). The atrioturbinal is more rostrocaudally elongated from the early stage in *R. affinis* and *H. gentilis* and from the mid stage in *R. pusillus* (**Figures 5A,A′,F,F′, 6A,A′**). In addition, the atrioturbinal developed toward the rostrocaudal direction starting with the late stage (**Figures 5, 6**). While the atrioturbinal was visible, the marginoturbinal was partly visible in our scans. Hence, the marginoturbinal cannot be reconstructed. The contrast between the thick cartilage and surrounding soft tissue was not clear enough to identify the boundary.

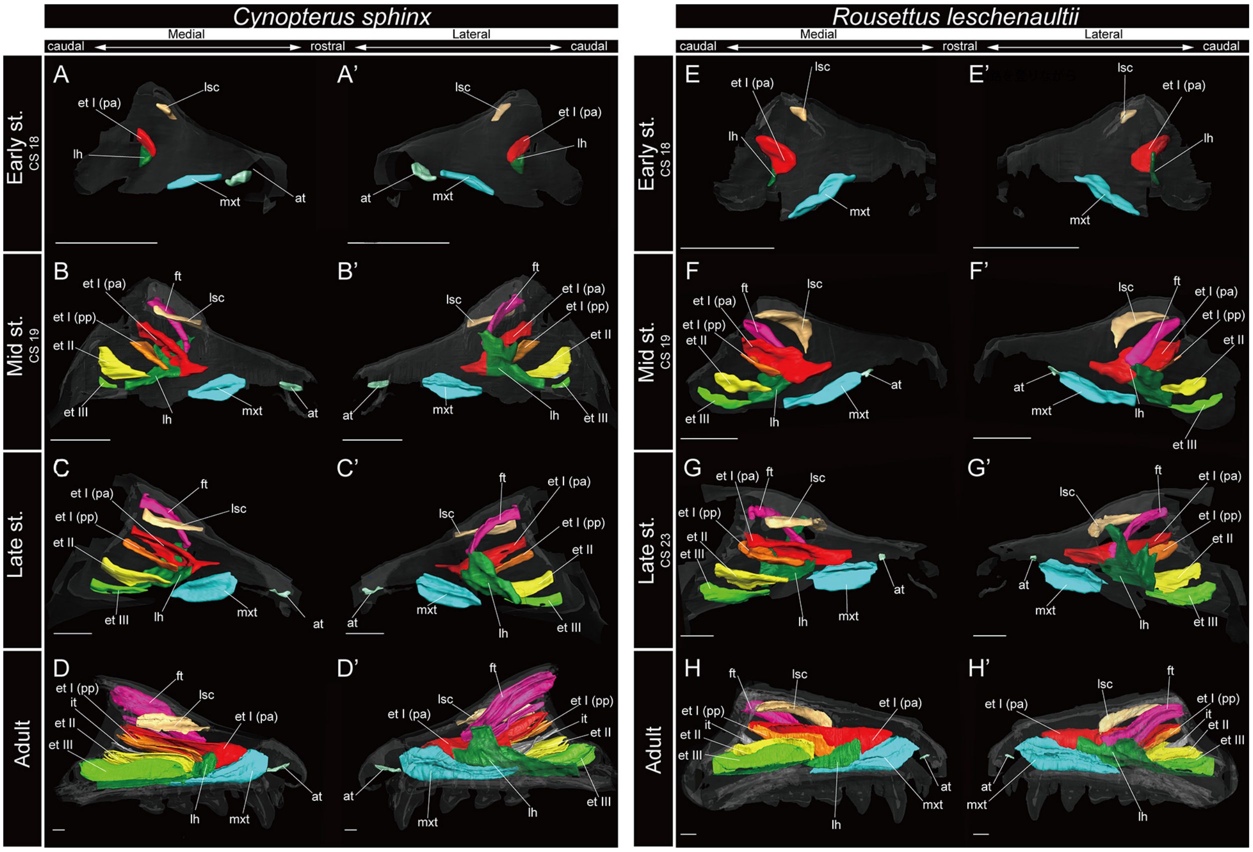


FIGURE 4 | 3D reconstructions of the developing turbinals in *Cynopterus sphinx* and *Rousettus leschenaultii*. **(A–D′)** *C. sphinx*; **(E–H′)** *R. leschenaultii*; **(A,A′,E,E′)** early stage fetus; **(B,B′,F,F′)** mid stage fetus; **(C,C′,G,G′)** late stage fetus; **(D,D′,H,H′)** adult. **(A–D)** Medial view and **(A′–D′)** lateral view of *C. sphinx*; **(E–H)** medial view and **(E′–H′)** lateral view of *R. leschenaultii*. Scale bars, 1 mm. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pp), ethmoturbinal I pars posterior; et II–IV, ethmoturbinals II–IV; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; nt, nasoturbinal; lsc, lamina semicircularis.

Maxilloturbinal

In most specimens examined in this study, the maxilloturbinal was positioned caudally to the atrioturbinal. The maxilloturbinal was a rostrally positioned structure within the nasal cavity, and its ventral side folded inward. The maxilloturbinal enlarged as it developed in all outgroup species (**Figure 3** and **Supplementary Figure 2**). At the same time, it showed a double scroll in *S. scrofa* (**Figures 3C–D′** and **Supplementary Figures 2C,D**) and a single scroll in *S. murinus* and *F. catus* (**Figures 3A–B′,E–F′** and **Supplementary Figures 2A,B,E,F**). The maxilloturbinal of the outgroup fetus was cartilaginous.

Among bats, the developmental pattern of Pteropodidae resembled that in outgroup species. The maxilloturbinal of Pteropodidae was the largest among all turbinals and laminae starting in the early stage (**Figures 4A,A′,E,E′**). Beginning at the mid stage in *C. sphinx* and late stage in *R. leschenaultii*, the maxilloturbinal started branching (**Supplementary Figures 3B, 4C**). From the mid stage in *C. sphinx* and late stage in *R. leschenaultii*, the maxilloturbinal developed dorsal and ventral branches which were both laterally scrolled as in the late stage of *S. scrofa* (**Figures 3D,D′, 4B,B′,G,G′** and **Supplementary Figures 2D, 3B, 4C**). Also, the cartilaginous structure was replaced and ossified in the adult (**Supplementary Figures 3, 4**).

Within Rhinolophoidea, all species presented similar maxilloturbinal morphologies. The maxilloturbinal enlarged and only partially ossified in the early to late stages (**Figures 5, 6** and **Supplementary Figures 5–8**). The maxilloturbinal also fused with the lamina horizontalis caudally such that it occurred in the early stage in *Rhinolopus* and *A. stoliczkanus* (**Figures 5A,A′,E,E′, 6E,E′**) and in the mid stage in *H. gentilis* (**Figures 6B,B′**). Nonetheless, the maxilloturbinal was reduced compared with other turbinals and laminae after the late stage in all species belonging to Rhinolophoidea (**Figures 5, 6** and **Supplementary Figures 5–9**). In addition, only the caudal side of the maxilloturbinal was ossified in the adult, and the rostral side remained cartilaginous (**Supplementary Figure 9**).

The maxilloturbinal of *M. siligorensis* and *V. sinensis* was rostrocaudally elongated and lateromedially narrow (**Figure 7**). It was slim and rod-shaped from the early to late stages (**Figures 7A–C′,E–G′**). Unlike the outgroup and Pteropodidae, the maxilloturbinal did not branch, and it showed a single scroll ventrally as it developed from the late stage to adult (**Supplementary Figures 2–4, 10, 11**). It extended lateromedially,

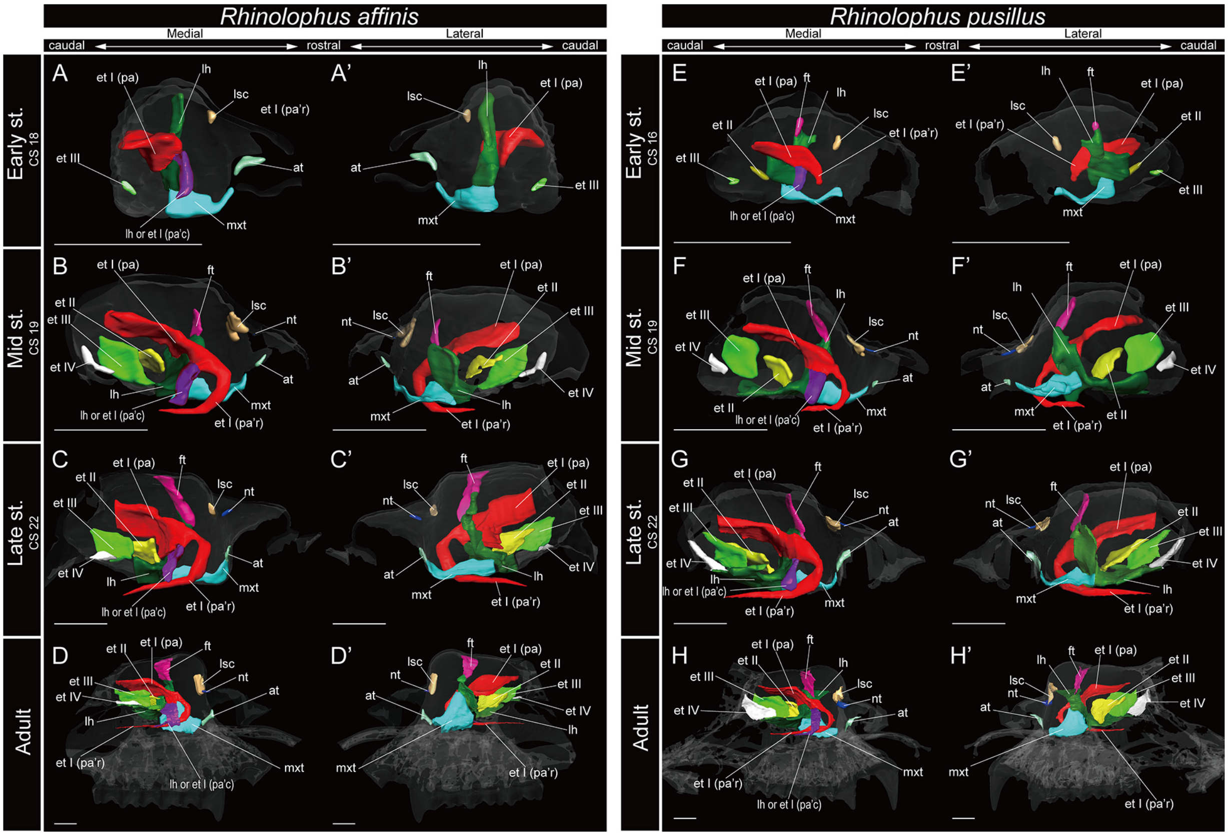


FIGURE 5 | 3D reconstructions of the developing turbinals in *Rhinolophus affinis* and *R. pusillus*. **(A–D′)** *R. affinis*; **(E–H′)** *R. pusillus*; **(A,A′,E,E′)** early stage fetus; **(B,B′,F,F′)** mid stage fetus; **(C,C′,G,G′)** late stage fetus; **(D,D′,H,H′)** adult. **(A–D)** Medial view and **(A′–D′)** lateral view of *R. affinis*; **(E–H)** medial view and **(E′–H′)** lateral view of *R. pusillus*. Scale bars, 1 mm. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pa'c), caudal part of ethmoturbinal I pars anterior; et I (pa'r), rostral part of ethmoturbinal I pars anterior; et II–IV, ethmoturbinals II–IV; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; nt, nasoturbinal; lsc, lamina semicircularis.

becoming a plate-like structure where it attached to the inner lateral nasal wall in the adult (**Figures 7D,D′,H,H′** and **Supplementary Figures 10, 11**).

Nasoturbinal

We observed the nasoturbinal in mid and late stages of the outgroup species. The nasoturbinal slightly projected ventrally from the nasal wall (**Figure 3**). It did not show any scrolling, and it extended rostrocaudally and was observed near the naris in both mid and late stages in all outgroups (**Figure 3** and **Supplementary Figure 2**). The length of the nasoturbinal varied among the outgroup species such that the nasoturbinal of *S. murinus* and *F. catus* was rostrocaudally longer than *S. scrofa* (**Figure 3**). The nasoturbinal of *S. murinus* and *F. catus* elongated rostrocaudally such that its length was comparable to that of the maxilloturbinal (**Figures 3A–B′,E–F′**).

In Pteropodidae, the nasoturbinal was absent during prenatal developmental stages (**Figure 4** and **Supplementary Figures 3, 4**). In the adult, a slight projection was observed dorsally to the nasal cavity near the naris (**Supplementary Figures 3, 4**).

The nasoturbinal of *Rhinolophus* and *H. gentilis* from the mid stage and of *A. stoliczkanus* from the late stage consisted of a tiny cartilaginous structure, continuing with the lamina semicircularis (**Figures 5B–D′,F–H′**, **6B–D′,G–H′** and **Supplementary Figures 5–8**).

The nasoturbinal of *M. siligorensis* and *V. sinensis* was much more well-developed than other chiropteran species and projected slightly ventrally (**Figure 7**). It did not show any scrolling and extended rostrocaudally beyond the atrioturbinal–maxilloturbinal contact. While it formed a short rod-like structure in the early and mid stages (**Figures 7A–B′,E–F′** and **Supplementary Figures 10, 11**), in the late stage and adult, it formed a long rod-like structure rostrocaudally (**Figures 7C–D′,G–H′** and **Supplementary Figures 10, 11**).

Lamina Semicircularis

The lamina semicircularis was observed in all outgroup species (**Figure 3**). This lamina extended from the inner wall in the central region of the nasal capsule toward the lateromedial, dorsoventral, and caudorostral directions. It expanded transversally on the dorsal side of the nasal cavity

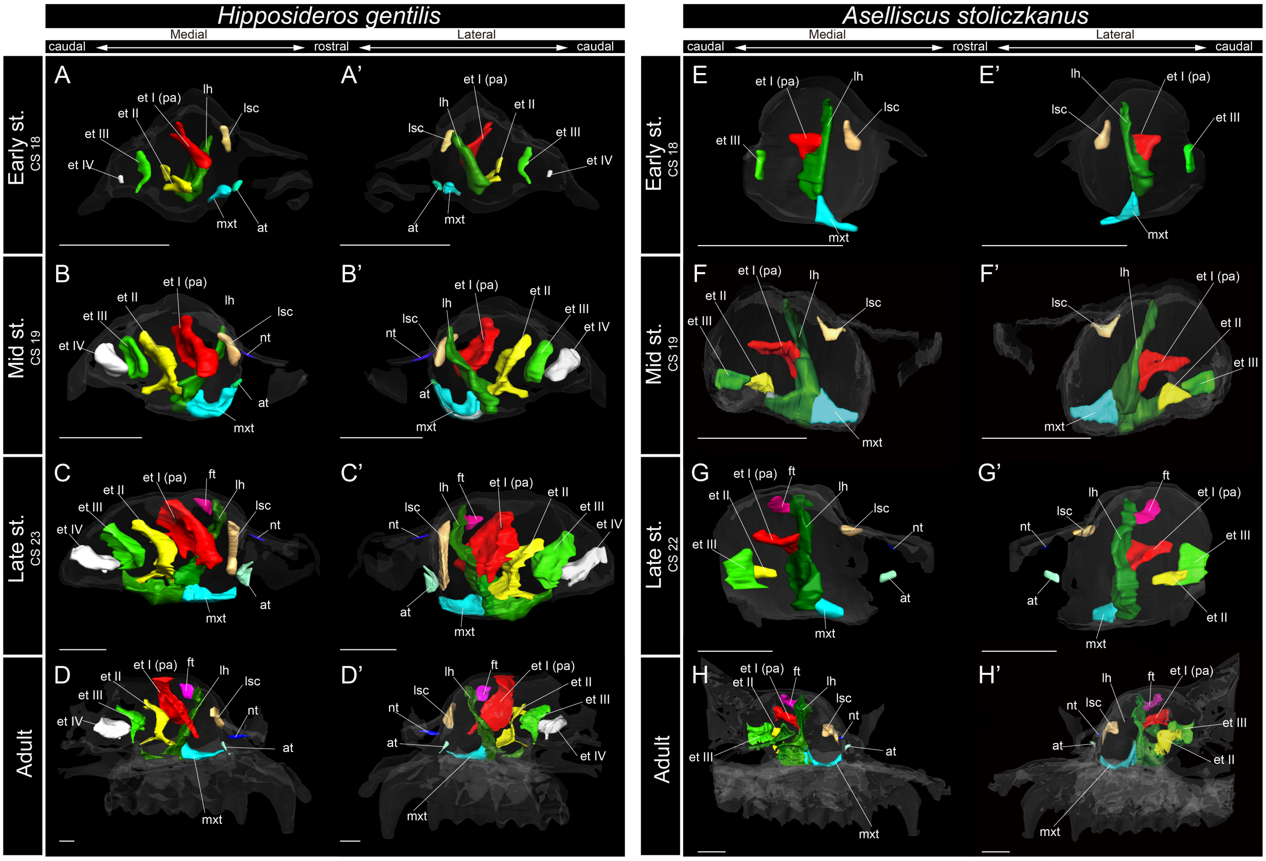


FIGURE 6 | 3D reconstructions of the developing turbinals in *Hipposideros gentilis* and *Aselliscus stoliczkanus*. **(A–D′)** *H. gentilis*; **(E–H′)** *A. stoliczkanus*; **(A,A′,E,E′)** early stage fetus; **(B,B′,F,F′)** mid stage fetus; **(C,C′,G,G′)** late stage fetus; **(D,D′,H,H′)** adult. **(A–D)** Medial view and **(A′–D′)** lateral view of *H. gentilis*; **(E–H)** medial view and **(E′–H′)** lateral view of *A. stoliczkanus*. Scale bars, 1 mm. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pp), ethmoturbinal I pars posterior; et II–IV, ethmoturbinals II–IV; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; nt, nasoturbinal; lsc, lamina semicircularis.

as it developed from the mid stage to the late (**Figure 3** and **Supplementary Figure 2**).

Among bats, the lamina semicircularis of *C. sphinx* and *R. leschenaultii* showed an almost identical developmental pattern as that of the outgroup species in terms of the transverse expansion. The lamina semicircularis was observed in the early stage for both species (**Figures 4A,A′,E,E′** and **Supplementary Figures 3, 4**). The lamina semicircularis extended in the caudorostral direction rather than the dorsoventral direction which is different from the observation in the outgroup. While the lamina semicircularis of *C. sphinx* and *R. leschenaultii* was not as large as that of the outgroup species, it was as large as the frontoturbinal and the ethmoturbinal III from the mid stage (**Figures 4B–D′,F–H′**).

The developmental pattern of the lamina semicircularis was similar between Rhinolophoidea species (**Figures 5**, **6** and **Supplementary Figures 5–8**). The lamina semicircularis was observed in all fetal stages and in the adult of Rhinolophoidea (**Figures 5**, **6** and **Supplementary Figures 5–8**). The lamina semicircularis of *H. gentilis* was the largest among all Rhinolophoidea (**Figures 5**, **6** and **Supplementary Figures 5–8**). It projected ventrally from the inner wall of the nasal capsule starting in the early stage (**Figures 6A,A′** and **Supplementary Figure 7A**). The lamina semicircularis then elongated toward the lateromedial and dorsoventral directions from the inner wall of the nasal capsule in the mid to late stages (**Figures 6B–C′** and **Supplementary Figures 7B,C**). In the late stage, the lateromedially and dorsoventrally extended lamina semicircularis formed a wall that separated the anterior and the middle region of the nasal capsule (**Figures 6C,C′** and **Supplementary Figure 7C**). In the adult, the lamina semicircularis was partly ossified, but not scrolled in all Rhinolophoidea (**Figures 5**, **6** and **Supplementary Figures 5–8**).

In the adult of *M. siligorensis* and *V. sinensis*, the lamina semicircularis was not observed in the caudal region of the nasal cavity, which was surrounded by the maxilla and palatine (**Figures 7D,D′,H,H′** and **Supplementary Figures 10, 11**). The nasoturbinal was found in most of the dorsal region of the external nasal cartilage (rostral side of the nasal cavity), but the laminar structure was not seen in the caudal side of the nasoturbinal (**Figures 7D,D′,H,H′** and **Supplementary Figures 10, 11**). While the nasoturbinal was observed, the laminar structure was also not observed in the caudal side of the nasoturbinal in any fetal stages (**Figures 7A–C′,E–G′** and

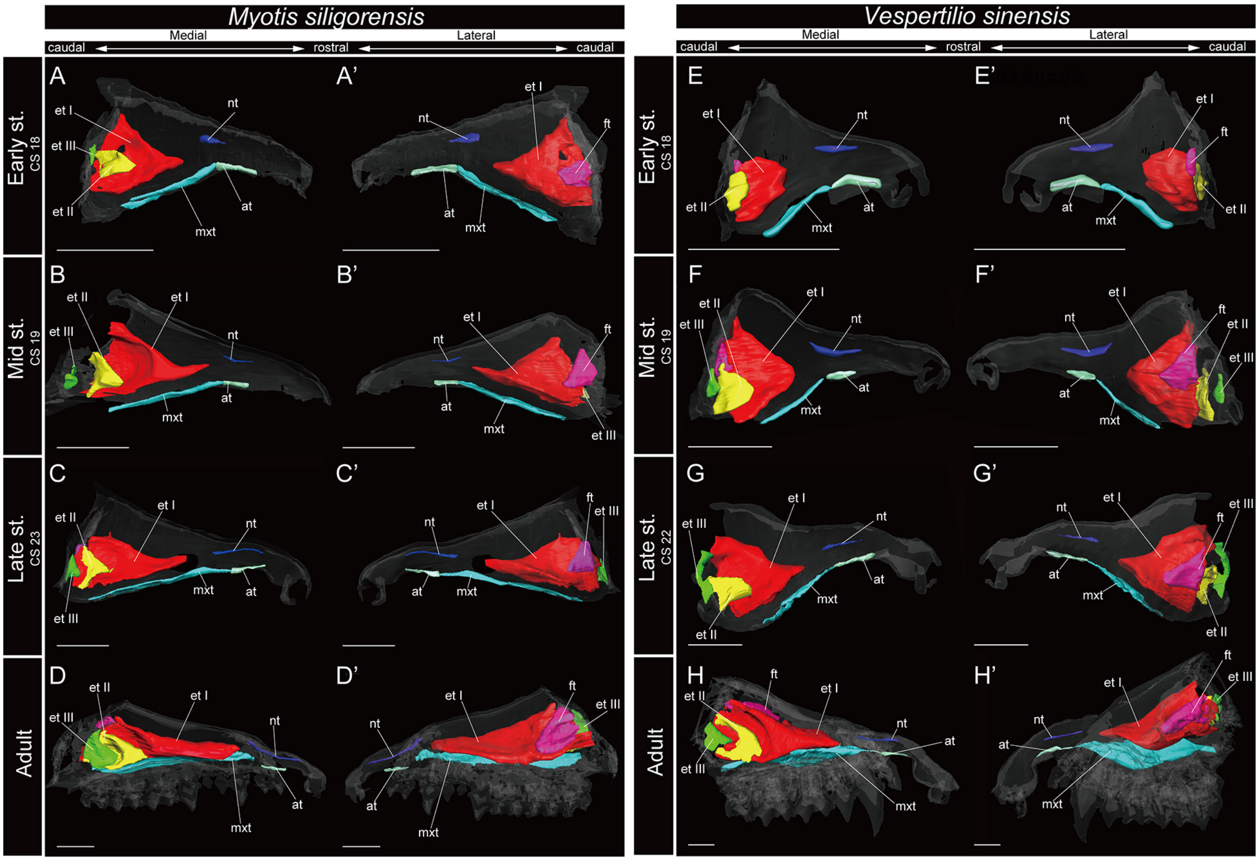


FIGURE 7 | 3D reconstructions of the developing turbinals in *Myotis siligorensis* and *Vespertilio sinensis*. **(A–D′)** *M. siligorensis*; **(E–H′)** *V. sinensis*; **(A,A′,E,E′)** early stage fetus; **(B,B′,F,F′)** mid stage fetus; **(C,C′,G,G′)** late stage fetus; **(D,D′,H,H′)** adult. **(A–D)** Medial view and **(A′–D′)** lateral view of *M. siligorensis*; **(E–H)** medial view and **(E′–H′)** lateral view of *V. sinensis*. Scale bars, 1 mm. at, atrioturbinal; et I (pa), ethmoturbinal I pars anterior; et I (pp), ethmoturbinal I pars posterior; et II–III, ethmoturbinal II–III; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; nt, nasoturbinal; lsc, lamina semicircularis.

Supplementary Figures 10, 11). Hence, the lamina semicircularis was not formed in both *M. siligorensis* and *V. sinensis*.

Lamina Horizontalis

The lamina horizontalis horizontally separates the nasopharyngeal duct and the ethmoturbinal recess, which includes several turbinals in the outgroup species (**Figure 3** and **Supplementary Figure 2**). The lamina horizontalis of the outgroup species extended from the inner wall of the nasal capsule toward the lateromedial, rostrocaudal, and caudorostral directions. In the mid and late stages, the dorsal side of the lamina horizontalis fused with the ventral side of the ethmoturbinal I pars anterior and the ethmoturbinal I pars posterior except for the mid stage of *S. scrofa* (**Figure 3** and **Supplementary Figure 2**).

Among bats, *C. sphinx* and *R. leschenaultii* showed a developmental pattern similar to the outgroup species, whereas the developmental patterns of Rhinolophoidea and Yangochiroptera were different from those of the outgroup species. Similar to the mid stage of *S. scrofa*, the lamina horizontalis projected from the inner wall of the nasal capsule in the early stage of *C. sphinx* and *R. leschenaultii*, and dorsally, it fused with the ventral side of the ethmoturbinal I pars anterior (**Figures 3C,C′, 4A,A′,E,E′** and **Supplementary Figures 2C, 3A, 4A**). The lamina horizontalis likely extended toward the lateromedial, rostrocaudal, and caudorostral directions and fused with the ventromedial part of the ethmoturbinal I pars posterior and the ethmoturbinal II from the early to the mid stages (**Figures 4A–B′,E–F′** and **Supplementary Figures 3, 4**). The lamina horizontalis enlarged lateromedially and elongated rostrocaudally from the mid stage to adult (**Figures 4B–D′,F–H′** and **Supplementary Figures 3, 4**). In all fetal stages, it was cartilaginous in both species, but ossified in the adult (**Supplementary Figures 3, 4**).

The lamina horizontalis of Rhinolophoidea showed a different developmental pattern compared with that of all outgroups and Pteropodidae (**Figures 3–6**). In all members of Rhinolophoidea, the lamina horizontalis that projected from the inner wall of the nasal capsule extended toward the lateromedial and dorsoventral directions from the early stage to adult. The lamina horizontalis formed a wall perpendicular to the rostrocaudal plane (**Figures 5, 6** and **Supplementary Figures 5–8**).

Among Rhinolophoidea, the presumptive developmental pattern of the lamina horizontalis of *Rhinolophus* was somewhat dissimilar to that of *H. gentilis* or *A. stoliczkanus*, such that the lamina horizontalis scrolled ventrolaterally beginning in the early stage (**Figures 5A,A′,E,E′** and **Supplementary Figures 5A, 6A**). The apex of the lamina horizontalis projected inward in the early stage and then gradually extended ventrally and medially from the mid stage to the adult (**Figure 5** and **Supplementary Figures 5, 6**). In the adult, its apex extended caudally into the nasopharyngeal duct (**Supplementary Figures 5, 6**). When observed medially from the sagittal plane, it formed a hairpin-shaped structure (**Figures 5D,D′,H,H′**). When seen in coronal sections, the lamina horizontalis extended from the dorsolateral side of the nasal wall toward the ventromedial side. At the medial portion of the nasal capsule, the ethmoturbinal I pars anterior is positioned dorsally from the lamina horizontalis from the early stage of *Rhinolophus* (**Supplementary Figures 5, 6**). The ventromedial part of the lamina horizontalis bent lateralward from the point in which the ethmoturbinal I pars anterior extended from the late stage to adult in *R. affinis* and from mid sage to adult of *R. pusillus* (**Supplementary Figures 5, 6**). At the caudal portion of the nasal cavity of adult *Rhinolophus*, the ventral edge of the ventromedial part of the lamina horizontalis was round and appeared inside of the nasopharyngeal duct (**Supplementary Figures 5, 6**). Nonetheless, this hairpin-shaped structure of the lamina horizontalis may be the ethmoturbinal I pars anterior (**Figure 5**). The structure which is most probably the lamina horizontalis was ossified in the adult (**Supplementary Figures 4D, 5D, 9**).

The lamina horizontalis was not observed in *M. siligorensis* and *V. sinensis* in any fetal stages or the adult, unlike all outgroup species as well as Pteropodidae and Rhinolophoidea (**Figure 7** and **Supplementary Figures 10, 11**).

Ethmoturbinal I Pars Anterior

The ethmoturbinal I pars anterior of the outgroup species projected toward the lateromedial and caudorostral directions from the inner wall of the nasal capsule already in the mid stage, fusing with the dorsal side of the lamina horizontalis (**Figure 3** and **Supplementary Figure 2**). It was the largest turbinal among all ethmoturbinals from the mid to late stages in all outgroups (**Figure 3**). The ethmoturbinal I pars anterior of *S. scrofa* was less developed than that of *F. catus* and *S. murinus* in the mid stage (**Figures 3A,A′,C,C′,E,E′**).

The ethmoturbinal I pars anterior of Pteropodidae showed the same developmental pattern as that of the outgroup species (**Figures 3**, **4**). The developmental pattern from the early to mid stages, in particular, resembled that from the mid to the late stages of *S. scrofa* (**Figures 3C–D′**, **4A–B′,E–F′**). The ethmoturbinal I pars anterior of Pteropodidae fused ventrally with the lamina horizontalis and projected toward the lateromedial direction from the inner wall of the nasal capsule (**Figures 4A–B′,E–F′**). The ethmoturbinal I pars anterior extended rostrally from the early to mid stages, and in the adult, it reached as far as the dorsal border of the maxilloturbinal (**Figures 4A–B′,D,D′,E–F′,H,H′**). The ethmoturbinal I pars anterior of Pteropodidae was the largest turbinal among all ethmoturbinals from the early stage to the adult (**Figure 4**). The ethmoturbinal I pars anterior was ossified in the adult (**Supplementary Figures 4, 5**).

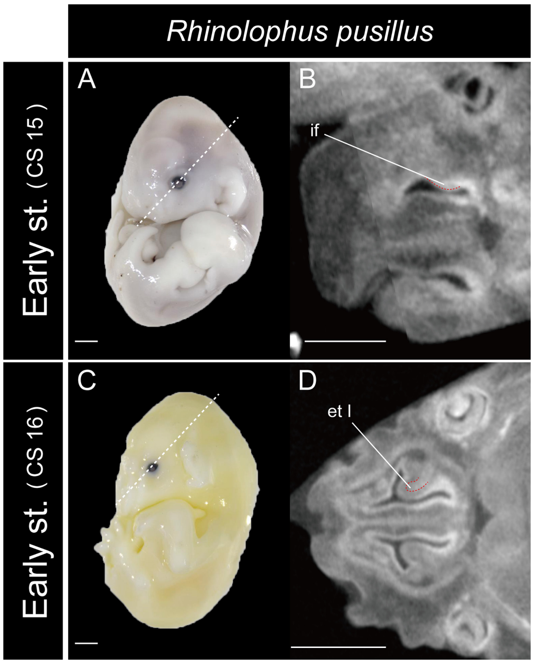


FIGURE 8 | The onset of the turbinate projections in *Rhinolophus pusillus*. Broken lines indicate the location of section. Scale bar, 1 mm. **(A)** Early stage (CS 15) fetus embryo; **(B)** early stage fetus (CS 15) cross section; **(C)** early stage fetus (CS 16) embryo; **(D)** early stage fetus (CS 16) cross section. et I, ethmoturbinal I; if, initial fold.

In *Rhinolophus*, the developmental pattern of the ethmoturbinal I pars anterior differed from that of all other species and formed a distinctive structure (**Figure 5**). The early stage (CS 15) of *R. pusillus* was the smallest of all samples examined in this study (**Figures 8A,B**). In the early stage (CS 15) of *R. pusillus*, the cartilage of the nasal capsule was obscured in the scan, but an initial fold was observed in the inner wall of the nasal capsule (**Figure 8B**). The cartilage which we believe as the ethmoturbinal I pars anterior of the late phase of early stage (CS 16) was likely embedding within this initial fold of the early stage (CS 15) in *R. pusillus* (**Figures 8C,D**).

The fusion of the ethmoturbinal I pars anterior with the lamina horizontalis extended toward the lateromedial direction from the inner wall in the early stage of *Rhinolophus*, like the mid stage of *S. scrofa*, and the early stage of Pteropodidae (**Figures 3–5**, and **Supplementary Figures 2–6**). The alternative interpretation is that the lamina horizontalis did not form the hairpin-shaped structure in medial side. Consequently, the ethmoturbinal I pars anterior split into rostral and caudal parts from the early stage (**Figures 5A,A′,E,E′** and **Supplementary Figures 5, 6**). In this case, the rostral part of the ethmoturbinal I pars anterior extended rostrally and curved ventrally in the mid stage of *R. affinis* and the early stage of *R. pusillus* (**Figures 5B,B′,E,E′**). Subsequently, it entered the nasopharyngeal duct, extending caudally in the mid stage of *R. affinis* and the late stage of *R. pusillus* (**Supplementary Figures 5B, 6C**). When observed medially from the sagittal plane, it formed a hairpin-shaped structure (**Figures 5B,B′,E,E′**). The apex of the rostral part of the ethmoturbinal I pars anterior extended caudally also from the mid stage (**Figures 5B,B′,E,E′**). It formed the freestanding structure in the adult (**Figure 5**).

The caudal part of the ethmoturbinal I pars anterior, which can be identified as a part of the lamina horizontalis in the alternative interpretation, scrolled medially in the mid stage of *R. affinis* and *R. pusillus* (**Figures 5B,B′,F,F′** and **Supplementary Figures 5, 6**). Subsequently, the apex of the caudal part of the ethmoturbinal I pars anterior that projected inward gradually extended ventrally and medially until the late stage (**Figure 5**). In the adult, this apex extended caudally and became another freestanding structure in the nasopharyngeal duct (**Figures 5D,D′,H,H′** and **Supplementary Figures 5, 6**). From the sagittal plane, it also formed a hairpin-shaped structure medially (**Figures 5D,D′,H,H′**). When seen in coronal sections, ethmoturbinal I pars anterior was present at the rostral portion of the nasal capsule. At the medial portion of the nasal capsule, the rostral and caudal part of ethmoturbinal I pars anterior was branched off from the lamina horizontalis. The rostral part extended toward the dorsomedial side, and the caudal part toward the ventromedial side from the early stage to adult (**Supplementary Figures 5, 6**). At the caudal portion of the nasal cavity of adult, the ventral edge of the caudal part of ethmoturbinal I pars anterior was round and appeared inside of the nasopharyngeal duct medially to the rostral part of ethmoturbinal I pars anterior (**Supplementary Figures 5, 6**).

In the early stage of *H. gentilis* and *A. stoliczkanus*, the ventral side of the ethmoturbinal I pars anterior fused at the dorsal side of the dorsoventrally elongated lamina horizontalis and projected from the inner wall of the nasal capsule (**Figures 6A,A′,E,E′** and **Supplementary Figures 7, 8**). The ethmoturbinal I pars anterior was large after the early stage, and it developed dorsally and caudally to the lamina horizontalis in *H. gentilis* and *A. stoliczkanus*, respectively (**Figures 6A,A′,E,E′**). Still, it did not elongate rostrally like that of *S. murinus* and Pteropodidae (**Figures 3A–B′**, **4**, **5**). The ethmoturbinal I of Rhinolophoidea was ossified in the adult (**Supplementary Figures 5–8**).

Similar to the outgroup species and other bats, the ethmoturbinal I of *M. siligorensis* and *V. sinensis* was the largest turbinal from the early stage to adult (**Figure 7**). On the other hand, it differed from the outgroup species and Pteropodidae in that it did not split into pars anterior and pars posterior (**Figures 3**, **4**, **7** and **Supplementary Figures 2–4, 10, 11**). The ventral side of the rostral end of ethmoturbinal I formed a horizontal plate-like process from the mid stage to adult (**Figures 7B–D′,F–H′** and **Supplementary Figures 10, 11**). In addition, since the lamina horizontalis was absent in all stages, ethmoturbinal I projected solely from the inner wall of the nasal capsule (**Figure 7** and **Supplementary Figures 10, 11**). While it enlarged dorsoventrally from the early to late stages (**Figures 7A–C′,E–G′** and **Supplementary Figures 10, 11**), ethmoturbinal I extended toward the caudorostral direction from the late stage to adult (**Figures 7C–D′,G–H′** and **Supplementary Figures 10, 11**). Ethmoturbinal I of Yangochiroptera was ossified in the adult (**Supplementary Figures 10, 11**).

Ethmoturbinal I Pars Posterior

The ethmoturbinal I pars posterior was observed from the ventral side of the ethmoturbinal I pars anterior starting in the mid stage of *S. murinus* and *F. catus* and in the late stage of *S. scrofa* (**Figures 3A,A′,D,D′,E,E′** and **Supplementary Figures 2A,D,E**).

Ethmoturbinal I pars posterior of Pteropodidae was located at the same place as that of the outgroup species (**Figures 3, 4**). It was absent in the early stage (**Figures 4A,A′,E,E′** and **Supplementary Figures 3, 4**) but appeared starting in the mid stage (**Figures 4B,B′,F,F′** and **Supplementary Figures 3, 4**).

In contrast to the outgroup species and Pteropodidae, ethmoturbinal I pars posterior was absent from the early stage to adult in all members of Rhinolophoidea, *M. siligorensis*, and *V. sinensis* (**Figures 5–7**).

Turbinals in the Ethmoturbinal Recess

Ethmoturbinals II and III projected from the inner wall of the nasal capsule starting in the mid stage in *S. murinus* and *F. catus* (**Figures 3A,A′,E,E′** and **Supplementary Figures 2A,E**). In *S. scrofa*, ethmoturbinal II was observed in the mid stage, while ethmoturbinal III was only present in the late stage (**Figures 3C–D′** and **Supplementary Figures 2C,D**). In the outgroup species, ethmoturbinal II was the second largest among all ethmoturbinals after ethmoturbinal I pars anterior except for *S. murinus* (**Figure 3**).

In Pteropodidae, ethmoturbinal II and ethmoturbinal III were absent in the early stage; however, they projected from the inner wall of the nasal capsule after the mid stage (**Figure 4** and **Supplementary Figures 3, 4**). In terms of size, ethmoturbinal II was larger compared with ethmoturbinal III in Pteropodidae in the mid and late stages (**Figures 4B–C′,F–G′**). Ethmoturbinal III was larger than ethmoturbinal I pars posterior from the late stage to adult (**Figures 4C–D′,G–H′**).

In *R. affinis*, ethmoturbinal II appeared from the inner wall of the nasal capsule in the mid stage, while in *R. pusillus*, it was observed in a similar position in the early stage (**Figures 5B,B′,E,E′** and **Supplementary Figures 5B, 6A**). In both species, ethmoturbinal III projected from the inner wall of the nasal capsule in the early stage (**Figures 5A,A′,E,E′** and **Supplementary Figures 5**, **6**). In *R. pusillus*, the size of ethmoturbinal II and ethmoturbinal III was mostly comparable in the early stage (**Figures 5E,E′**). Ethmoturbinal III was larger compared with ethmoturbinal II in the mid stage (**Figures 5B,B′,F,F′**). Also, ethmoturbinal IV arose in the mid stage of *Rhinolophus* (**Figures 5B,B′,F,F′** and **Supplementary Figures 5**, **6**). Ethmoturbinal III was the largest, and ethmoturbinals II and IV were of the same size among these three turbinals from the late stage *Rhinolophus* (**Figures 5C,C′,G,G′**).

In *H. gentilis*, ethmoturbinals II, III, and IV projected from the inner wall of the nasal capsule in the early stage (**Figures 6A,A′**). Ethmoturbinal III was large in the early stage (**Figures 6A,A′**). Moreover, in *H. gentilis*, ethmoturbinal II was larger than ethmoturbinal III in the mid stage; however, the size became similar from the late stage (**Figures 6B–D′**). In *A. stoliczkanus*, ethmoturbinal III appeared in the early stage (**Figures 6E,E′**).

Then, in the mid stage, ethmoturbinal II appeared rostrally to ethmoturbinal III (**Figures 6F,F′**). Ethmoturbinal IV was not observed from the early stage to adult (**Figures 6H,H′**).

Ethmoturbinal II appeared from the inner wall of the nasal capsule in the early stage of *M. siligorensis* and *V. sinensis*; however, after the mid stage, ethmoturbinal II shifted laterally, approaching ethmoturbinal I during late ontogeny as the nasal capsule enlarged (**Figures 7B–D′,F–H′** and **Supplementary Figures 10, 11**). The caudal part of ethmoturbinal II curved and fused with the laterally positioned ethmoturbinal I from the late stage to adult (**Figures 7C–D′,G–H′** and **Supplementary Figures 10, 11**). Both ethmoturbinal II and ethmoturbinal III extended toward the caudorostral and the ventrodorsal directions, forming a triangular shape from the mid stage (**Figures 7B–D′,F–H′**).

DISCUSSION

Turbinate Ontogeny and Homology in Laurasiatheria and Pteropodidae

We found that the turbinate structures are principally comparable between Laurasiatheria and Pteropodidae. The fetuses of outgroup species, *S. murinus*, *S. scrofa*, and *F. catus*, together with the fetuses of Pteropodidae, appear to share a ventrally positioned and enlarged maxilloturbinal, which is the largest among all turbinals (**Figures 3**, **4** and **Supplementary Figures 2–4**). Ethmoturbinal I is the largest to develop among ethmoturbinals, splitting into pars anterior and pars posterior in all outgroup species (**Figures 3A–B′,D–F′**). In our study, adult specimens of outgroup species were not examined. Previous studies showed the ventrally positioned, branched off, and enlarged maxilloturbinal and ethmoturbinal I pars anterior and pars posterior in the adult *S. murinus* (Kuramoto, 1980), *S. scrofa* (Paulli, 1900a), and *F. catus* (Moore, 1981). These characteristics were also seen in the adult Pteropodidae studied here. Furthermore, the lamina horizontalis of the outgroup species and Pteropodidae divides the nasopharyngeal duct and the ethmoturbinal recess that includes several turbinals (**Figures 3**, **4** and **Supplementary Figures 2–4**). The enlarged ethmoturbinal recess with a vast space formed dorsally which is filled with the olfactory mucosa suggests that Pteropodidae are equipped with high olfactory ability (Negus, 1958; Craven et al., 2007, 2010; Eiting et al., 2014b). The developmental structure and position of the turbinal and lamina of the outgroup species studied here are congruent with the generality seen among previously reported non-volant mammals (Paulli, 1900a,b,c; Voit, 1909; Moore, 1981; Smith and Rossie, 2008; Van Valkenburgh et al., 2014).

Some authors mention the possibility that the unique turbinate morphology of *Rhinolophus* (which was also seen in our study) is related to echolocation (Curtis and Simmons, 2017; Curtis et al., 2020). Given this, and the fact that *Rhinolophus* belongs to Yinpterochiroptera as do Pteropodidae, we expected to see similar morphology in echolocating *Rousettus* prior to our experiment (Feldhamer et al., 2007). However, such behavioral differences (*Rousettus* is known for using tongue clicks for echolocation, while *Cynopterus* does not engage in such behavior) are not reflected in the turbinate structures of *R. leschenaultii* and *C. sphinx* (**Figure 4** and **Supplementary Figures 3, 4**).

Based on adult specimens, Allen (1882) and Paulli (1900c) suggested that Pteropodidae [*Cynopterus*, *Epomophorus gambianus*, *Pteropus giganteus*, *Pteropus* sp., *Rousettus* (*Cyonycteris*)] have “four endoturbinals.” Given Allen’s (1882) schematic for the turbinal of *E. gambianus*, we assume that he incorrectly split the true ethmoturbinal I into “endoturbinal I” and “endoturbinal II” (**Supplementary Figure 12**). Allen (1882) drew other schematics for the turbinal of non-volant mammals, in which the author identified ethmoturbinal I pars anterior and pars posterior as the “endoturbinal I and lobule.” This suggests that the misidentification of the turbinal of *E. gambianus* reported by Allen (1882) was clearly not caused by the difference in nomenclature. Only tentative inferences can be made, as Paulli (1900c) provided no schematic for the turbinals of Pteropodidae, but we assume Paulli labeled the lamina semicircularis as “endoturbinal I.” Thus, there are “four endoturbinals” for the Pteropodidae species in Paulli’s view. Giannini et al. (2012) identified the turbinals based on a histological section of one fetal stage of *Pteropus* sp. and a CT-scanned image of the adult *P. lylei*. They pointed out that there is one additional “endoturbinal” in both studies of Allen (1882) and Paulli (1900c) compared with their observation (**Figures 4C–D′,G–H′** and **Supplementary Figure 12**). Our observation on the turbinals of *C. sphinx* and *R. leschenaultii* shows that ethmoturbinal III projects from the inner nasal wall from the mid stage (**Figures 4B–D′,F–H′** and **Supplementary Figures 3, 4**). Following this, no more ethmoturbinals are formed (**Figure 4** and **Supplementary Figures 3, 4**). Allen’s study on *E. gambianus* (with our assumption that ethmoturbinal I splits into two parts; **Supplementary Figure 12**), Giannini’s study on *Pteropus*, and our study on *C. sphinx* and *R. leschenaultii* indicate that the number of ethmoturbinals is three in Pteropodidae (**Figure 4** and **Supplementary Figures 3, 4**).

Giannini et al. (2012) compared *Pteropus* with non-bat mammals and claimed that the turbinate element composition in *Pteropus* is comparable with that of non-human primates (Smith and Rossie, 2006) and rodents (Paulli, 1900c) in terms of the reduced number of turbinals. They also concluded that the number of frontoturbinals in *Pteropus* differs from that of the hedgehog *Erinaceus* (Paulli, 1900c) and the marsupial *Monodelphis* (Rowe et al., 2005; Macrini, 2014) such that *Pteropus* has one while the hedgehog and marsupial have two frontoturbinals. The turbinate structure of *Pteropus* is both similar and different from that of carnivorans and ungulates which have increasing number and complexity (Stößel et al., 2010) in ethmoturbinals, frontoturbinals, and interturbinals (Giannini et al., 2012). Carnivorans have three ethmoturbinals like Pteropodidae, which are similar to our results (Paulli, 1900a,b,c; Wagner and Ruf, 2019, 2020) (**Supplementary Table 1**). However, Paulli reported that the maximum of the frontoturbinal and interturbinal combined (ectoturbinal in Paulli) is five to ten (only *Meles*) (Paulli, 1900c). This is different to Pteropodidae with one frontoturbinal and one interturbinal. The number of ethmoturbinal of ungulates varies such that

Capra has three and *Dicotyles labiatus* has seven. Moreover, the number of frontoturbinal and interturbinal largely varies with seven in *Tragulus javanicus* and 31 in *Equus caballus* (Paulli, 1900c; Moore, 1981). *Capra* showed the same number of ethmoturbinal as Pteropodidae; however, the number of ethmoturbinal and frontoturbinal and interturbinal combined in ungulates is generally larger than that of Pteropodidae. Nonetheless, Giannini et al. (2012) pointed out that the turbinate homology of elements in *Pteropus* can be traced without difficulty based on the position of turbinals in canids (*Vulpes vulpes*). Even though the number varies for certain turbinals (caudal ethmoturbinal, interturbinal, and frontoturbinal) among non-volant Laurasiatheria and Pteropodidae, the turbinal element composition of pteropodids is easily traceable from that of our non-volant laurasiatherians. Therefore, we agree with Giannini et al. (2012) that the turbinate element composition of Pteropodidae is rather similar to that of other mammals (even for species with complex turbinals).

Yangochiroptera

The lamina horizontalis is not formed throughout ontogeny in the studied Vespertilionidae (**Figure 7** and **Supplementary Figures 10, 11**). Consequently, there is no clear separation between the nasopharyngeal duct and the ethmoturbinal recess. This anatomical setting suggests that these species are not specialized to keep the inspired air within the nasal cavity, allowing for better odorant sorption (Negus, 1958; Adams, 1972; Craven et al., 2010; Eiting et al., 2014b). Based on the nasal cavity structure, we suggest that the studied Vespertilionidae members are less capable of catching odorants compared with the outgroup mammals and Pteropodidae which have the lamina horizontalis and an independent space of the ethmoturbinal recess.

Fetal turbinals of *P. pipistrellus*, *V. murinus* (Grosser, 1900), *M. schreibersii* (Fawcett, 1919; De Beer, 1937), and *M. myotis* (Frick, 1954) were studied previously using histological sections. The figures given by Fawcett (1919); De Beer (1937), and Frick (1954) did not show the lamina, which separates the nasopharyngeal duct and the ethmoturbinal recess that includes several turbinals in these Vespertilionidae species. The lamina horizontalis separates the nasopharyngeal duct and the ethmoturbinal recess in all stages in the outgroup species and after mid stage in Pteropodidae (**Figures 3**, **4** and **Supplementary Figures 2–4**). Our study is consistent with the observation of Fawcett (1919); De Beer (1937), and Frick (1954). In the early stage of *M. siligorensis* and *V. sinensis*, ethmoturbinal II is located on the inner wall of the nasal capsule (**Figures 7A,A′,E,E′** and **Supplementary Figures 10, 11**), but the caudal part of ethmoturbinal II curves and fuses with the laterally positioned ethmoturbinal I from the late stage to adult (**Figures 7C–D′,G–H′** and **Supplementary Figures 10, 11**). As identified by Fawcett (1919) and De Beer (1937) in *M. schreibersii* and Frick (1954) in *M. myotis*, our study confirms that ethmoturbinal II is positioned medially to ethmoturbinal I in all fetal stages and adult in *M. siligorensis* and *V. sinensis*.

The present study did not cover members of Phyllostomidae, Emballonuridae, Molossidae, and Nycteridae, which also belong to Yangochiroptera (Teeling et al., 2002, 2003, 2005). Although most bats of Yangochiroptera emit sonar from their oral apparatus, Phyllostomidae and Nycteridae are characterized by emitting sonar from the naris (Jones and Teeling, 2006; Feng et al., 2012). As for the turbinals of Nycteridae, Allen (1882) is the only study that identifies turbinals in this family, in which the author described the adult *Nycteris thebaica*. He stated that *N. thebaica* has two endoturbinals (endoturbinal I has a lobule) and one ectoturbinal. Moreover, he claimed that it has a nasoturbinal that is larger than the endoturbinal. Regarding Phyllostomidae, Bhatnagar and Kallen (1974a) studied *Artibeus jamaicensis*; Kämper and Schmidt (1977) studied *Artibeus lituratus*, *Carollia perspicillata*, *Glossophaga soricina*, *Phyllostomus discolor*, and *Sturnira lilium*; and Yohe et al. (2018) studied *A. jamaicensis*, *Brachyphylla pumila*, *Erophylla bombifrons*, and *Phyllonycteris poeyi* of Phyllostomidae. Principally, the members of Phyllostomidae reportedly have seven turbinals. Although details of the identification varies among studies, all these studies agree that in Phyllostomidae the ethmoturbinal recess within the nasal cavity is separated with the nasopharyngeal duct rostrally by the ethmoturbinal and caudally by the lamina (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Yohe et al., 2018). Presenting coronal sections of the ethmoturbinal recess in three bats (*Anoura geoffroyi*, *S. lilium*, *Uroderma bilobatum*), Eiting et al. (2014a) showed that the lamina that separates the nasopharyngeal duct and the ethmoturbinal recess is well-developed. However, these studies are all based on adult species, and fetal information of Phyllostomidae is still largely lacking (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Eiting et al., 2014a; Yohe et al., 2018).

Compared with the patterns reported for Phyllostomidae, we recognize that *M. siligorensis* and *V. sinensis* do not have the lamina horizontalis, which is the lamina separating the nasopharyngeal duct and ethmoturbinal recess (**Figure 7**). This was particularly obvious in the caudal region of the nasal cavity (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Eiting et al., 2014a; Yohe et al., 2018). Further observation of the nasal development of Phyllostomidae, Mollossidae, Emballonuridae, and Nycteridae is required to clarify the whole picture of turbinate homology within Yangochiroptera.

Rhinolophoidea

The rostral part of the lamina horizontalis extends dorsally in Rhinolophoidea from the early stage (**Figures 5**, **6A–A′,E–E′** and **Supplementary Figures 5–8**). The lamina horizontalis pushes the ethmoturbinal recess back toward the caudal direction, resulting in a small ethmoturbinal recess. As the size of the ethmoturbinal recess likely relates to olfactory ability (Negus, 1958; Adams, 1972; Craven et al., 2010; Eiting et al., 2014b), it is likely that Rhinolophoidea may have a reduced olfactory ability compared with the outgroup species and Pteropodidae.

Rhinolophoidea are undoubtedly the most disputed and problematic taxon among bats regarding its turbinate homology. Members of the superfamily Rhinolophoidea, which include Rhinolophidae, Hipposideridae, Megadermatidae, Craseonycteridae, and Rhinopomatidae, emit echolocation

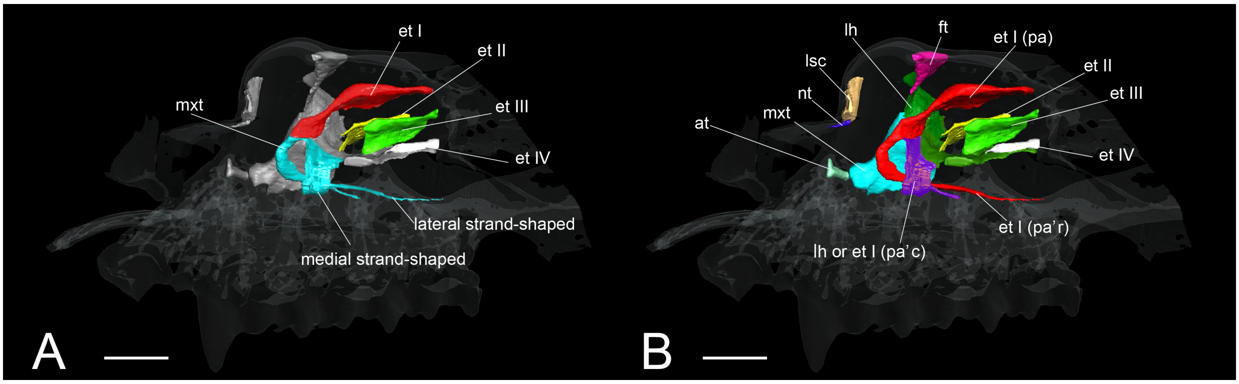


FIGURE 9 | The identification of turbinals and laminas in *Rhinolophus affinis* by Curtis and Simmons (2017) and this study. **(A)** Descriptions by Curtis and Simmons (2017) with strand-shaped maxilloturinal; **(B)** revised descriptions by this study. Scale bar, 1 mm. et I (pa), ethmoturbinal I pars anterior; et I (pa'c), caudal part of ethmoturbinal I pars anterior; et I (pa'r), rostral part of ethmoturbinal I pars anterior; et II–IV, ethmoturbinals II–IV; ft, frontoturbinal; mxt, maxilloturbinal; lh, lamina horizontalis; lsc, lamina semicircularis.

pulses from the naris (Csorba et al., 2003; Teeling et al., 2005; Li et al., 2007). The relationship between this behavior and their turbinal anatomy has not been understood (Curtis and Simmons, 2017; Curtis et al., 2020). Until today, very few have studied the turbinals of Rhinolophoidea, and there is some confusion in the literature regarding their turbinal homology. Studying *R. ferrumequinum* and *Rhinolophus hipposideros*, Grosser (1900) described the maxilloturbinal as a freestanding structure within the nasopharyngeal duct (**Supplementary Figure 13**). Following Grosser's identification, recent studies on Rhinolophidae identified the freestanding structure in the nasopharyngeal duct as the maxilloturbinal (Curtis and Simmons, 2017; Curtis et al., 2020) (**Figure 9**). Curtis and Simmons (2017) and Curtis et al. (2020) reported that the maxilloturbinal forms two strand-shaped structures that project rostrally and enter the nasopharyngeal duct, referring to these structures as the lateral and medial strand of the maxilloturbinals (Curtis and Simmons, 2017; Curtis et al., 2020).

In our view, their "maxilloturbinal" is not the true maxilloturbinal (**Figure 9**). Our observations on fetal and adult specimens in *Rhinolophus* show that the true maxilloturbinal is reduced, and its caudal end may be partially ossified (**Supplementary Figure 9**). We cannot be certain that it is ossified as we observed using the iodine-based solution. Observing the histological sections provided by Curtis et al. (2020), we can see that the structure we believe to be the maxilloturbinal (Curtis et al., 2020 also identifies a part of this as the maxilloturbinal) is unossified. Other Rhinolophoidea members, *H. gentilis*, and *A. stoliczkanus* also show similar maxilloturbinal like *Rhinolophus* with reduced and partially ossified caudal end even in the adult (**Supplementary Figures 5–8**). Curtis et al. (2020) also show the histological sections of *Hipposideros lankadiva*, and what we see as maxilloturbinal is unossified. In *Hipposideros*, the identification of the maxilloturbinal is congruent among their and our studies. Nonetheless, they do not present a 3D reconstruction; thus, we cannot be definite.

The anatomical definition of the maxilloturbinal is described in Maier's therian bauplan for the nasal capsule (**Figure 1**) (Maier, 1993a). Maier has demonstrated that the maxilloturbinal is continuous with the atrioturbinal, observing Primates, Prosimii (*Daubentonia madagascariensis*, *Galagoides demidoff*) and Platyrrhini (*Pithecia monachus*, *Saimiri sciureus*), and Scandentia (*Ptilocercus lowii*, *Tupaia belangeri*) (Maier, 1980, 2000; Maier and Ruf, 2014). The fact that the maxilloturbinal is continuous with the atrioturbinal appears to be the common pattern for therian mammals (Maier, 1993a). Following Maier's bauplan, we identify the structure that is continuous with the atrioturbinal as the maxilloturbinal. The maxilloturbinal in Rhinolophoidea extends from the atrioturbinal in the early stage in *R. affinis* and *H. gentilis*, the mid stage in *R. pusillus*, and the late stage in *A. stoliczkanus* (**Figures 5A,A′,F,F′, 6A,A′,G,G′** and **Supplementary Figures 5–8**).

So, what are the lateral and medial strands of the "maxilloturbinal" that project into the nasopharyngeal duct shown by Grosser (1900); Curtis and Simmons (2017), and Curtis et al. (2020)? In our view, their lateral strand of the maxilloturbinal is probably the ethmoturbinal pars anterior. In mammals, the first projection within the nasal capsule becomes ethmoturbinal I (Smith and Rossie, 2008). Our results show that presumptive initial fold first appears within the nasal capsule in the fetus (CS 15) of *R. pusillus* (**Figures 8A,B**). Then, the cartilaginous structure of ethmoturbinal I projects toward the initial fold in the fetus (CS 16) (**Figures 8B,D**).

The projecting cartilaginous structure of ethmoturbinal I pars anterior extends rostrally and turns slightly ventrally within the nasal capsule in the early stage of *R. pusillus* (**Figures 5E,E′**). The ethmoturbinal I pars anterior forms a hairpin-shaped structure with a distinctive curve during the mid stage to adult in *Rhinolophus* such that its tip extends caudally (**Figures 5B–D′,E–H′**). The position of the formed ethmoturbinal I pars anterior partly matches that of the lateral strand of the "maxilloturbinal" reported by Curtis and Simmons (2017) (**Figure 9**).

There are two possible interpretations regarding the "medial strand of the maxilloturbinal" of *Rhinolophus*. We cannot confirm whether it is part of the lamina horizontalis or part of the ethmoturbinal I pars anterior at this point. This is because in the region where the medial hairpin-shaped turbinal is formed, the boundary between the lamina horizontalis and the ethmoturbinal I is indistinctive, making the identification of the medial hairpin-shaped turbinate structure difficult. De Beer (1937) identified the lamina horizontalis as part of ethmoturbinal I. Our rationale is based on the observation of the topologies of turbinals during development, and further identification is unreasonable.

If the medial hairpin-shaped turbinate structure is the lamina horizontalis, it elongates medially in the early stage and extends ventrally and turns laterally from the mid to late stages (**Figures 5A–C′,E–G′**). In the adult, its apex extends caudally (**Figures 5D,D′,H,H′**). A comprehensive study on Rhinolophoidea is required to test this scenario.

If the medial hairpin-shaped turbinate structure is a part of ethmoturbinal I pars anterior, ethmoturbinal I pars anterior of *Rhinolophus* splits into rostral and the caudal parts. The rostral part of ethmoturbinal I pars anterior is the “lateral strand of the maxilloturbinals” in Curtis and Simmons (2017), and the caudal part is their “medial strand of the maxilloturbinals.” The caudal part of ethmoturbinal I pars anterior extends medially, then turns ventrally and continues toward the caudal direction from the early stage in *Rhinolophus* (**Figures 5A,A′,E,E′**). We found it as the medial hairpin-shaped turbinate structure within the nasopharyngeal duct in adult (**Figures 5D,D′,H,H′**, **9B** and **Supplementary Figures 5, 6**).

We conclude that it is unlikely that the lateral hairpin-shaped turbinate structure is the ethmoturbinal I pars anterior and the medial hairpin-shaped turbinate structure is ethmoturbinal I pars posterior in Rhinolophidae. Ethmoturbinal I pars posterior is formed medially to ethmoturbinal I pars anterior and dorsally to the lamina horizontalis in Pteropodidae and in various non-bat mammals (Allen, 1882; Voit, 1909; Smith and Rossie, 2008; Maier and Ruf, 2014; Ruf et al., 2015; Ruf, 2020). Nonetheless, the medial hairpin-shaped turbinate structure is formed ventrally to the lamina horizontalis in *Rhinolophus*. The position is too far apart from where we would expect (where we can locate the ethmoturbinal I pars posterior in non-bat mammals).

Lateral and medial hairpin-shaped turbinate structures have not been reported from any previous study in other Rhinolophoidea including Hipposideridae (*Aselliscus tricuspidatus*, *Coleops frithii*, *Hipposideros armiger*, *Hipposideros fulvus*, *Hipposideros pratti*, and *Hipposideros speoris*), *Rhinopoma*, *Macroderma gigas*, *M. lyra*, and *Rhinonycteris aurantia* (Nelson et al., 2007; Smith et al., 2012; Curtis and Simmons, 2017; Curtis et al., 2020). However, as previous studies did not observe prenatal specimens and did not investigate both soft and hard parts of the turbinals, they did not rule out the possibility that the hairpin-shaped turbinate structure may be present in Hipposideridae (Curtis and Simmons, 2017). We found that ethmoturbinal I (the turbinal that forms the hairpin-shaped turbinal in Rhinolophidae) of Hipposideridae, *H. gentilis* and *A. stoliczkanus*, projects from the inner wall of the nasal capsule in the early stage and then enlarges toward the rostral direction after the early stage (**Figure 6**), confirming that these species do not form the hairpin-shaped turbinal. The lateral and medial hairpin-shaped turbinate structures that we identify as ethmoturbinal I pars anterior or the lamina horizontalis are certainly unique to Rhinolophidae.

Character Evolution of Turbinals in Bats

The possible evolutionary scenario of the turbinal architecture in bats is summarized in **Figure 10**. Although several Eocene fossil bats including *Onychonycteris finneyi* have been reported (Simmons et al., 2008), their nasal capsule and nasal cavity are difficult to observe due to fossilization. Thus, we can only infer ancestral traits from extant species. Giannini et al. (2012) pointed out that the turbinate morphology of Pteropodidae is highly comparable to that of non-volant laurasiatherians. We found that Rhinolophoidea are characterized by the vertically stranding lamina horizontalis and the rostrally cartilaginous and caudally ossified maxilloturbinal (**Figure 10**). The lamina horizontalis does not separate the nasopharyngeal duct and the ethmoturbinal recess in our limited Yangochiroptera sample. Given these observations, the turbinate morphology seen in Pteropodidae is unlikely to be a result of convergence, and we assume that the bat common ancestor most probably had a turbinate morphology comparable to Pteropodidae and non-volant Laurasiatheria.

In both Pteropodidae and Rhinolophoidea, the lamina horizontalis separates the nasopharyngeal duct and ethmoturbinal recess. On the other hand, the maxilloturbinal is developed in Pteropodidae while it is reduced in Rhinolophoidea, and the orientation of the lamina horizontalis is horizontal in Pteropodidae while it is vertical in Rhinolophoidea. Ethmoturbinal I pars posterior is lost in Rhinolophoidea. It is reported that *R. aurantia* and *Triaenops persicus* have a small number of ethmoturbinals, while *M. lyra* has an increased number of ethomoturbinals (I–VII) (Smith et al., 2012; Curtis and Simmons, 2017). Thus, the diversification of ethmoturbinal numbers appears to characterize the Rhinolophoidea lineage.

Our results and previous literature on *M. schreibersii* (Fawcett, 1919; De Beer, 1937) and *M. myotis* (Frick, 1954) suggest that Yangochiroptera have lost ethmoturbinal I pars posterior (**Figure 10**). As noted earlier, ethmoturbinal I pars posterior is also lost in Rhinolophoidea. If the condition of the bat common ancestor was the same as Pteropodidae and non-volant laurasiatherians, the loss of ethmoturbinal I pars posterior in Yangochiroptera and Rhinolophoidea has occurred independently. Studies by Fawcett (1919); De Beer (1937), and Frick (1954), and ours suggest that oral-emitting Yangochiroptera (*M. schreibersii*, *M. myotis*, *M. siligorensis*, and *V. sinensis*) lack the lamina horizontalis, which separates the nasopharyngeal duct and ethmoturbinal recess in Yinpterochiroptera. On the other hand, nasal-emitting Yangochiroptera such as Phyllostomidae possess the structure that separates the nasopharyngeal duct and ethmoturbinal recess (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Yohe et al., 2018). This raises a new question: whether the structure that separates the nasopharyngeal duct and ethmoturbinal recess in Yinpterochiroptera and Phyllostomidae is homologous or not. In Yinpterochiroptera, the separation between the nasopharyngeal duct and ethmoturbinal recess in the rostral part of the nasal cavity is achieved by the lamina horizontalis. In contrast, the separation is achieved

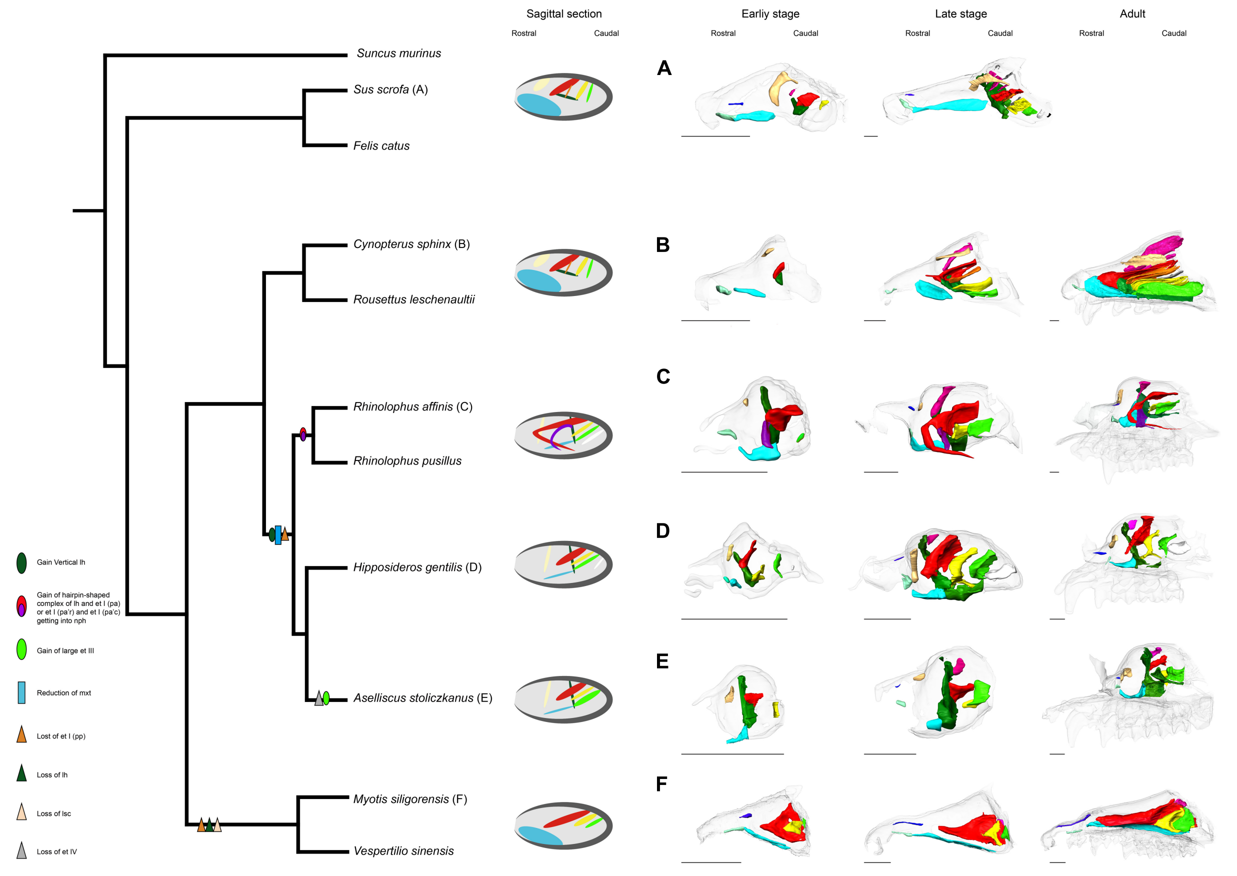

Sagittal section
Rostral
Caudal
Earliy stage
Rostral
Caudal
Late stage
Rostral
Caudal
Adult
Rostral
Caudal
Suncus murinus
Sus scrofa (A)
Felis catus
Cynopterus sphinx (B)
Rousettus leschenaultii
Rhinolophus affinis (C)
Rhinolophus pusillus
Hipposideros gentilis (D)
Aselliscus stoliczkanus (E)
Myotis siligorensis (F)
Vespertilio sinensis
A
B
C
D
E
F
Gain Vertical lh
Gain of hairpin-shaped complex of lh and et I (pa) or et I (pa'r) and et I (pa'c) getting into nph
Gain of large et III
Reduction of mxt
Lost of et I (pp)
Loss of lh
Loss of lsc
Loss of et IV

FIGURE 10 | Inferred evolutionary history of the nasal structures. **(A)** *S. scrofa*; **(B)** *C. sphinx*; **(C)** *R. affinis*; **(D)** *H. gentilis*; **(E)** *A. stoliczkanus*; **(F)** *M. siligorensis*. Note that the inference for Yangochiroptera is based only on Vespertilionidae. ethmoturbinal I pars anterior (red); ethmoturbinal I pars posterior (orange); ethmoturbinal II (yellow); ethmoturbinal III (light green); ethmoturbinal IV (white); lamina horizontalis (green); lamina horizontalis or caudal part of ethmoturbinal I pars anterior (purple); lamina semicircularis (flesh color); maxilloturbinal (light blue).

by the complex structure of ethmoturbinals in Phyllostomidae (Bhatnagar and Kallen, 1974a; Kämper and Schmidt, 1977; Yohe et al., 2018). It may be possible that the separation occurred secondarily in Phyllostomidae after the lamina horizontalis was lost in the common ancestor of Yangochiroptera, although this needs to be examined through observations on prenatal specimens of Phyllostomidae.

Using fetal specimens of various bat species, we have described the detailed 3D development of the nasal turbinals in bats and mostly resolved the confused homology of turbinals, though there are still some questions regarding the lamina horizontalis in Yangochiroptera and rostral turbinate structures of bats. Our study emphasizes the importance of studying prenatal anatomy and observing 3D structures of turbinals to address its homology problems. However, our study did not include members of Yangochiroptera from the New World, such as Phyllostomidae, Mollossidae, Emballonuridae, and Nycteridae. Currently, whether laryngeal echolocation has a single origin in bats or evolved multiple times independently is still disputed (Teeling et al., 2005; Nojiri et al., in press). The character states of the nasal turbinals of the common ancestor and how nasal turbinals have evolved with the evolution of laryngeal echolocation are still unknown. We envision future studies on the prenatal anatomy of bats to clarify the picture of their turbinal evolution and also solve the remaining problems associated with the homology of this complex structure among Laurasiatheria.

CONCLUSION

Using diceCT imaging, we comparatively described the 3D prenatal development of the nasal cavity in eight bat species of Yangochiroptera and Yinpterochiroptera and three species of non-volant Laurasiatheria. By observing multiple stages of nasal development, we solved the confused turbinate homology among bats and clarified the evolutionary history of the nasal turbinals in bats. We found that the strand-shaped structure of the "maxilloturbinal" of Rhinolophidae in Grosser (1900); Curtis and Simmons (2017), and Curtis et al. (2020) is not the true maxilloturbinal. We conclude that the "maxilloturbinal" of Grosser (1900); Curtis and Simmons (2017), and Curtis et al. (2020) is actually the complex of a part of the lamina horizontalis and ethmoturbinal I pars anterior or the rostral and caudal splitting parts of ethmoturbinal I pars anterior. The true maxilloturbinal is an undeveloped structure with a cartilaginous rostral part and an ossified caudal part, even in the adult. We found that the turbinate structures are principally comparable between Laurasiatheria and Pteropodidae, suggesting that Pteropodidae retain the basic condition of Laurasiatheria.

Pteropodidae exhibit an enlarged ethmoturbinal recess similar to non-volant mammals with a keen olfactory sense. The ethmoturbinal recess is significantly smaller in Rhinolophoidea compared with its closely related Pteropodidae, suggesting its reduced capability of olfaction. The lack of separation between the nasopharyngeal duct and ethmoturbinal recess in oral-emitting Yangochiroptera may indicate they are not well specialized for odorant deposition along the olfactory epithelium.

In both Pteropodidae and Rhinolophoidea, the lamina horizontalis separates the nasopharyngeal duct and ethmoturbinal recess. The maxilloturbinal is well developed in Pteropodidae, while it is reduced in Rhinolophoidea. The orientation of the lamina horizontalis is horizontal in Pteropodidae, while it is vertical in Rhinolophoidea. Rhinolophoidea are characterized by a well-developed vertically standing lamina horizontalis. It also acquired a rostrally cartilaginous and caudally ossified maxilloturbinal. The absence of ethmoturbinal I pars posterior in Yangochiroptera and Rhinolophoidea has occurred independently by convergent evolution. The separation of the nasopharyngeal duct and ethmoturbinal recess is found in Yinpterpchiroptera and Phyllostomidae, but not in oral-emitting Yangochiroptera. Whether the separation structure found in Yinpterpchiroptera and Phyllostomidae is homologous or it has evolved secondarily in Phyllostomidae should be tested in future studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available from the authors upon request.

ETHICS STATEMENT

The animal study was reviewed and approved by Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology.

AUTHOR CONTRIBUTIONS

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

FUNDING

This study was supported by JSPS #18H04816, #18H02492, #18K19359, #18KK0207, JRPs-LEAD with DFG to DK and Sasakawa Scientific Research grant 2020-5014 and JSPS # 18J14168 to KI.

ACKNOWLEDGMENTS

We thank Hiroki Higashiyama for technical support and Satoshi Suzuki for letting us observe samples under his care at the Kanagawa Prefectural Museum of Natural History.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.613545/full#supplementary-material

REFERENCES

Adams, D. R. (1972). Olfactory and non-olfactory epithelia in the nasal cavity of the mouse. *Peromyscus*. *Am. J. Anat.* 133, 37–49. doi: 10.1002/aja.1001330104

Allen, H. (1882). On a revision of the ethmoid bone in the Mammalia, with special reference to the description of this bone and of the sense of smelling in the Cheiroptera. *Bull. Mus. Comp. Zool.* 10, 135–171.

Bhatnagar, K. P., and Kallen, F. C. (1974a). Cribriform plate of ethmoid, olfactory bulb and olfactory acuity in forty species of bats. *J. Morphol.* 142, 71–89. doi: 10.1002/jmor.1051420104

Bhatnagar, K. P., and Kallen, F. C. (1974b). Morphology of the nasal cavities and associated structures in *Artibeus jamaicensis* and *Myotis lucifugus*. *Am. J. Anat.* 139, 167–189. doi: 10.1002/aja.1001390203

Bhatnagar, K. P., and Kallen, F. C. (1975). Quantitative observations on the nasal epithelia and olfactory innervation in bats. *Acta Anat.* 91, 272–282. doi: 10.1159/000144389

Brigham, R. M., Kalko, E. K. V., Jones, G., Parsons, S., and Limpens, H. (2004). *Bat Echolocation Research: Tools, Techniques and Analysis*. Austin, TX: Bat Conservation International.

Cooper, J. G., and Bhatnagar, K. P. (1976). Comparative anatomy of the vomeronasal organ complex in bats. *J. Anat.* 122, 571–601.

Craven, B. A., Neuberger, T., Paterson, E. G., Webb, A. G., Josephson, E. M., Morrison, E. E., et al. (2007). Reconstruction and morphometric analysis of the nasal airway of the dog (*Canis familiaris*) and implications regarding olfactory airflow. *Anat. Rec.* 290, 1325–1340. doi: 10.1002/ar.20592

Craven, B. A., Paterson, E. G., and Settles, G. S. (2010). The fluid dynamics of canine olfaction: unique nasal airflow patterns as an explanation of macrosmia. *J. R. Soc. Interface* 7, 933–943. doi: 10.1098/rsif.2009.0490

Cretekos, C. J., Weatherbee, S. D., Chen, C. H., Badwaik, N. K., Niswander, L., Behringer, R. R., et al. (2005). Embryonic staging system for the short-tailed fruit bat, *Carollia perspicillata*, a model organism for the mammalian order Chiroptera, based upon timed pregnancies in captive-bred animals. *Dev. Dyn.* 233, 721–738. doi: 10.1002/dvdy.20400

Csorba, G., Ujhelyi, P., and Thomas, N. (2003). *Horseshoe Bats of the World: (Chiroptera: Rhinolophidae)*. Bishop's Castle: Alana books.

Curtis, A. A., and Simmons, N. B. (2017). Unique turbinal morphology in horseshoe bats (Chiroptera: Rhinolophidae). *Anat. Rec.* 300, 309–325. doi: 10.1002/ar.23516

Curtis, A. A., Smith, T. D., Bhatnagar, K. P., Brown, A. M., and Simmons, N. B. (2020). Maxilloturbinal aids in nasophonation in horseshoe bats (Chiroptera: Rhinolophidae). *Anat. Rec.* 303, 110–128. doi: 10.1002/ar.23999

De Beer, G. (1937). *The Development of the Vertebrate Skull*. Oxford: Clarendon Press.

Deleon, V. B., and Smith, T. D. (2014). Mapping the nasal airways: using histology to enhance CT-based three-dimensional reconstruction in *Nycticebus*. *Anat. Rec.* 297, 2113–2120. doi: 10.1002/ar.23028

Dieulafe, L. (1906). Morphology and embryology of the nasal fossae of vertebrates. *Ann. Otol. Rhinol. Laryngol.* 15, 1–584.

Eiting, T. P., Smith, T. D., and Dumont, E. R. (2014a). Olfactory epithelium in the olfactory recess: a case study in new world leaf-nosed bats. *Anat. Rec.* 297, 2105–2112. doi: 10.1002/ar.23030

Eiting, T. P., Smith, T. D., Perot, J. B., and Dumont, E. R. (2014b). The role of the olfactory recess in olfactory airflow. *J. Exp. Biol.* 217, 1799–1803. doi: 10.1242/jeb.097402

Evans, H. E., and Sack, W. O. (1973). Prenatal development of domestic and laboratory mammals: growth curves, external features and selected references. *Anat. Histol. Embryol.* 2, 11–45. doi: 10.1111/j.1439-0264.1973.tb00253.x

Fawcett, M. D. (1919). The primordial cranium of *Miniopterus schreibersii* at the 17 millimetre total length stage. *J. Anat.* 53, 315–350.

Fehse, O. (1990). *Ein Beitrag zur Morphologie des Primordialcraniums von Rousettus aegyptiacus (Geoffroy 1810), (Megachiroptera, Mammalia)*. Tübingen: University of Tübingen.

Feldhamer, G. A., Drickamer, L. C., Vessey, S. H., Merritt, J. F., and Krajewski, C. (2007). *Mammalogy: Adaptation, Diversity, Ecology*, 3rd Edn. Baltimore: Johns Hopkins University Press.

Feng, L., Gao, L., Lu, H., and Müller, R. (2012). Nose leaf dynamics during pulse emission in horseshoe bats. *PLoS One* 7:e34685. doi: 10.1371/journal.pone.0034685

Frick, H. (1954). *Die Entwicklung und Morphologie des Chondrocraniums von Myotis Kaup*. Stuttgart: Georg Thieme Verlag.

Giannini, N. P., Macrini, T. E., Wible, J. R., Rowe, T. B., and Simmons, N. B. (2012). The internal nasal skeleton of the bat *Pteropus lylei* K. Andersen, 1908 (Chiroptera: Pteropodidae). *Ann. Carnegie Mus.* 81, 1–17. doi: 10.2992/007.081.0101

Giannini, N. P., Wible, J. R., and Simmons, N. B. (2006). On the cranial osteology of Chiroptera. I. *Pteropus* (Megachiroptera: Pteropodidae). *Bull. Am. Mus. Nat. Hist.* 295, 1–134. doi: 10.1206/0003-00902006295[0001:OTCOOC]2.0.CO;2

Gignac, P. M., and Kley, N. J. (2014). Iodine-enhanced micro-CT imaging: methodological refinements for the study of the soft-tissue anatomy of post-embryonic vertebrates. *J. Exp. Zool. Part B Mol. Dev. Evol.* 322, 166–176.

Gignac, P. M., Kley, N. J., Clarke, J. A., Colbert, M. W., Morhardt, A. C., Cerio, D., et al. (2016). Diffusible iodine-based contrast-enhanced computed tomography (diceCT): an emerging tool for rapid, high-resolution, 3-D imaging of metazoan soft tissues. *J. Anat.* 228, 889–909. doi: 10.1111/joa.12449

Göbbel, L. (2000). The external nasal cartilages in Chiroptera: significance for intraordinal relationships. *J. Mamm. Evol.* 7, 167–201. doi: 10.1023/A:1009436419637

Göbbel, L. (2002). Morphology of the external nose in *Hipposideros diadema* and *Lavia frons* with comments on its diversity and evolution among leaf-nosed Microchiroptera. *Cells Tissues Organs* 170, 39–60. doi: 10.1159/000047920

Gross, E. A., Swenberg, J. A., Fields, S., and Popp, J. A. (1982). Comparative morphometry of the nasal cavity in rats and mice. *J. Anat.* 135, 83–88.

Grosser, O. (1900). Zur Anatomie der Nasenhöhle und des Rachens der einheimischen Chiropteren. *Morphol. Jahrb.* 29, 1–77. doi: 10.1007/BF02301074

Hillenius, W. J. (1992). The evolution of nasal turbinates and mammalian endothermy. *Paleobiology* 18, 17–29. doi: 10.1017/S0094837300012197

Hutcheon, J. M., Kirsch, J. A. W., and Garland, T. (2002). A comparative analysis of brain size in relation to foraging ecology and phylogeny in the Chiroptera. *Brain Behav. Evol.* 60, 165–180. doi: 10.1159/000065938

Jones, G., and Teeling, E. C. (2006). The evolution of echolocation in bats. *Trends Ecol. Evol.* 21, 149–156. doi: 10.1016/j.tree.2006.01.001

Jurgens, J. D. (1962). *Contributions to the Descriptive and Comparative Anatomy of the Cranium of the Cape Fruit-Bat Rousettus aegyptiacus Leachi Smith*. [dissertation]. Stellenbosch: Zoological Institute, University of Stellenbosch.

Kämper, R., and Schmidt, U. (1977). Die Morphologie der Nasenhöhle bei einigen neotropischen Chiropteren. *Zoomorphologie* 87, 3–19. doi: 10.1007/BF02568739

Kuramoto, K. (1980). Morphological studies on the nasal turbinates of the musk shrew (*Suncus murinus riukiuanus*). *Nihon Juigaku Zasshi* 42, 377–380. doi: 10.1292/jvms1939.42.377

Li, G., Liang, B., Wang, Y., Zhao, H., Helgen, K. M., Lin, L., et al. (2007). Echolocation calls, diet, and phylogenetic relationships of Stoliczka's trident bat, *Aselliscus stoliczkanus* (Hipposideridae). *J. Mammal.* 88, 736–744. doi: 10.1644/06-MAMM-A-273R.1

Lozanoff, S., and Diewert, V. M. (1989). Developmental morphology of the solum nasi in the mouse lemur (*Microcebus murinus*). *J. Morphol.* 202, 409–424. doi: 10.1002/jmor.1052020308

Macrini, T. E. (2012). Comparative morphology of the internal nasal skeleton of adult marsupials based on X-ray computed tomography. *Bull. Am. Mus. Nat. Hist.* 365, 1–91. doi: 10.1206/365.1

Macrini, T. E. (2014). Development of the ethmoid in *Caluromys philander* (Didelphidae, Marsupialia) with a discussion on the homology of the turbinal elements in marsupials. *Anat. Rec.* 297, 2007–2017. doi: 10.1002/ar.23024

Maier, W. (1980). "Evolutionary biology of the new world monkeys and continental drift," in *Nasal Structures in Old and New World Primates*, eds R. L. Ciochon and A. B. Chiarelli (Boston, MA: Springer), 219–241.

Maier, W. (1993a). "Cranial morphology of the therian common ancestor, as suggested by the adaptations of neonate marsupials," in *Mammal Phylogeny*, eds F. S. Szalay, M. J. Novacek, and M. C. McKenna (New York, NY: Springer), 165–181.

Maier, W. (1993b). Zur evolutiven und funktionellen Morphologie des Gesichtsschädels der Primaten. *Z. Morph. Anthropol.* 79, 279–299.

Maier, W. (2000). "Ontogeny of the nasal capsule in cercopithecoids: a contribution to the comparative and evolutionary morphology of catarrhines," in *Old World Monkeys*, eds P. F. Whitehead and C. J. Jolly (Cambridge: Cambridge University Press), 99–132.

Maier, W., and Ruf, I. (2014). Morphology of the nasal capsule of Primates—with special reference to *Daubentonia* and *Homo*. *Anat. Rec.* 297, 1985–2006. doi: 10.1002/ar.23023

Martineau-Doizé, B., Caya, I., and Martineau, G.-P. (1992). Osteogenesis and growth of the nasal ventral conchae of the piglet. *J. Comp. Pathol.* 106, 323–331. doi: 10.1016/0021-9975(92)90018-p

Martinez, Q., Clavel, J., Esselstyn, J. A., Achmadi, A. S., Grohé, C., Pirot, N., et al. (2020). Convergent evolution of olfactory and thermoregulatory capacities in small amphibious mammals. *Proc. Natl. Acad. Sci. U.S.A.* 117, 8958–8965. doi: 10.1073/pnas.1917836117/-/DCSupplemental

Mead, J. G., and Fordyce, R. E. (2009). The therian skull: a lexicon with emphasis on the odontocetes. *Smithson. Contr. Zool.* 627, 1–249. doi: 10.5479/si.00810282.627

Moore, W. J. (1981). *The Mammalian Skull*. Cambridge: Cambridge University Press.

Negus, V. (1958). *Comparative Anatomy and Physiology of the Nose and Paranasal Sinuses*. Edinburgh: Livingstone.

Nelson, J. E., Christian, K. A., and Baudinette, R. V. (2007). Anatomy of the nasal passages of three species of Australian bats in relation to water loss. *J. Zool.* 55, 57–62. doi: 10.1071/ZO06101

Nojiri, T., Wilson, L. A. B., López-Aguirre, C., Tu, V. T., Kuratani, S., Ito, K., et al. (in press). Embryonic evidence uncovers convergent origins of laryngeal echolocation in bats. *Curr. Biol.* doi: 10.1016/j.cub.2020.12.043

Novacek, M. J. (1993). "Patterns of diversity in the mammalian skull," in *The Skull*, Vol. 2, eds J. Hanken and B. K. Hall (Chicago, IL: University of Chicago Press), 438–545.

Parker, W. K. (1885). *On the Structure and Development of the Skull in the Mammalia*. London: Royal Society of London.

Patterson, C. (1977). "Cartilage bones, dermal bones and membrane bones, or the exoskeleton versus the endoskeleton," in *Problems in Vertebrate Evolution*, eds S. M. Andrews, R. S. Miles, and A. D. Walker (London: Academic Press), 77–121.

Paulli, S. (1900a). Über die Pneumaticität des Schädels bei den Säugetieren. Eine morphologische Studie. I. Über den Bau des Siebbeins. Über die Morphologie des Siebbeins und die Pneumaticität bei den Monotremen und den Marsupialiern. *Gegenbaurs Morphol. Jahrb.* 28, 147–178.

Paulli, S. (1900b). Über die Pneumaticität des Schädels bei den Säugetieren. Eine morphologische Studie. II. Über die Morphologie des Siebbeins und die der Pneumaticität bei den Ungulaten und den Probosciden. *Gegenbaurs Morphol. Jahrb.* 28, 179–251.

Paulli, S. (1900c). Über die Pneumaticität des Schädels bei den Säugetieren. Eine morphologische Studie. III. Über die Morphologie des siebbeins und die der pneumaticität bei den Insectivoren, hyracoideen, Chiropteren, Carnivoren, pinnipedien, edentaten, Rodentiern, prosimiern und primaten, nebst einer zusammenfas-senden Übersicht über die Morphologie des siebbeins und die der pneumaticität des schädels bei den säugetieren. *Gegenbaurs Morphol. Jahrb.* 28, 483–564.

Reinbach, W. (1952a). Zur Entwicklung des Primordialcraniums von *Dasypus novemcinctus* Linné (*Tatusia novemcincta* Lesson) I. *Z. Morph. Anthropol.* 44, 375–444.

Reinbach, W. (1952b). Zur Entwicklung des Primordialcraniums von *Dasypus novemcinctus* Linné (*Tatusia novemcincta* Lesson) II. *Z. Morph. Anthropol.* 45, 1–72.

Rossie, J. B. (2006). Ontogeny and homology of the paranasal sinuses in Platyrrhini (Mammalia: Primates). *J. Morphol.* 267, 1–40. doi: 10.1002/jmor.10263

Rowe, T. B., Eiting, T. P., Macrini, T. E., and Ketcham, R. A. (2005). Organization of the olfactory and respiratory skeleton in the nose of the gray short-tailed opossum *Monodelphis domestica*. *J. Mamm. Evol.* 12, 303–336. doi: 10.1007/s10914-005-5731-5

Ruf, I. (2014). Comparative anatomy and systematic implications of the turbinal skeleton in Lagomorpha (Mammalia). *Anat. Rec.* 297, 2031–2046. doi: 10.1002/ar.23027

Ruf, I. (2020). Ontogenetic transformations of the ethmoidal region in muroidea (Rodentia, Mammalia): new insights from perinatal stages. *Vertebr. Zool.* 70, 383–415. doi: 10.26049/VZ70-3-2020-10

Ruf, I., Janßen, S., and Zeller, U. (2015). The ethmoidal region of the skull of *Ptilocercus lowii* (Ptilocercidae, Scandentia, Mammalia) – a contribution to the reconstruction of the cranial morphotype of primates. *Primate Biol.* 2, 89–110. doi: 10.5194/pb-2-89-2015

Scott, J. H. (1954). Heat regulating function of the nasal mucous membrane. *J. Laryngol. Otol.* 68, 308–317. doi: 10.1017/s0022215100049707

Simmons, N. B., and Cirranello, A. L. (2020). *Bat Species of the World: A Taxonomic and Geographic Database*. Available online at: https://batnames.org/ (accessed on 6 September 2020).

Simmons, N. B., Seymour, K. L., Habersetzer, J., and Gunnell, G. F. (2008). Primitive early eocene bat from wyoming and the evolution of flight and echolocation. *Nature* 451, 818–821. doi: 10.1038/nature06549

Smith, T. D., Curtis, A., Bhatnagar, K. P., and Santana, S. E. (2020). Fissures, folds, and scrolls: the ontogenetic basis for complexity of the nasal cavity in a fruit bat (*Rousettus leschenaultii*). *Anat. Rec*. doi: 10.1002/ar.24488. [Epub ahead of print].

Smith, T. D., Eiting, T. P., and Bhatnagar, K. P. (2012). A quantitative study of olfactory, non-olfactory, and vomeronasal epithelia in the nasal fossa of the bat *Megaderma lyra*. *J. Mamm. Evol.* 19, 27–41. doi: 10.1007/s10914-011-9178-6

Smith, T. D., Eiting, T. P., and Bhatnagar, K. P. (2015). "Anatomy of the nasal passages in mammals," in *Handbook of Olfaction and Gustation*, ed. R. L. Doty (Hoboken, NJ: John Wiley & Sons), 37–62.

Smith, T. D., and Rossie, J. B. (2006). "Primate olfaction: anatomy and evolution," in *Olfaction and the Brain: Window to the Mind*, eds W. Brewer, D. Castle, and C. Pantelis (Cambridge: Cambridge University Press), 135–166.

Smith, T. D., and Rossie, J. B. (2008). Nasal fossa of mouse and dwarf lemurs (Primates, Cheirogaleidae). *Anat. Rec.* 291, 895–915. doi: 10.1002/ar.20724

Sohn, J. H., Fukui, D., Nojiri, T., Minowa, K., Kimura, J., and Koyabu, D. (in press). Three-dimensional and histological observations on male genital organs of greater horseshoe bat, *Rhinolophus ferrumequinum*. *J. Mammal. Evol.* doi: 10.1007/s10914-020-09525-6

Springer, M. S., Teeling, E. C., Madsen, O., Stanhope, M. J., and De Jong, W. W. (2001). Integrated fossil and molecular data reconstruct bat echolocation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6241–6246. doi: 10.1073/pnas.111551998

Stößel, A., Junold, A., and Fischer, M. S. (2010). The morphology of the eutherian ethmoidal region and its implications for higher-order phylogeny. *J. Zool. Syst. Evol. Res* 48, 167–180. doi: 10.1111/j.1439-0469.2009.00560.x

Teeling, E. C., Madsen, O., Murphy, W. J., Springer, M. S., and O'Brien, S. J. (2003). Nuclear gene sequences confirm an ancient link between New Zealand's short-tailed bat and South American noctilionoid bats. *Mol. Phylogenet. Evol.* 28, 308–319. doi: 10.1016/S1055-7903(03)00117-9

Teeling, E. C., Madsen, O., Van Den Bussche, R. A., De Jong, W. W., Stanhope, M. J., and Springer, M. S. (2002). Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. *Proc. Natl. Acad. Sci. U.S.A* 99, 1431–1436. doi: 10.1073/pnas.022477199

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'Brien, S. J., and Murphy, W. J. (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307, 580–584. doi: 10.1126/science.1105113

Van Gilse, P. H. G. (1927). The development of the sphenoidal sinus in man and its homology in mammals. *J. Anat.* 61, 153–166.

Van Valkenburgh, B., Curtis, A. A., Samuels, J. X., Bird, D., Fulkerson, B., Meachen-Samuels, J., et al. (2011). Aquatic adaptations in the nose of carnivorans: evidence from the turbinates. *J. Anat.* 218, 298–310. doi: 10.1111/j.1469-7580.2010.01329.x

Van Valkenburgh, B., Smith, T. D., and Craven, B. A. (2014). Tour of a labyrinth: exploring the vertebrate nose. *Anat. Rec.* 297, 1975–1984. doi: 10.1002/ar.23021

Voit, M. (1909). Das Primordialcranium des Kaninchens unter Berüecksichtigung der Deckknochen. *Anat. Hefte* 38, 425–616.

Wagner, F., and Ruf, I. (2019). Who nose the borzoi? Turbinal skeleton in a dolichocephalic dog breed (*Canis lupus familiaris*). *Mamm. Biol.* 94, 106–119. doi: 10.13140/RG.2.2.24876.92800

Wagner, F., and Ruf, I. (2020). "Forever young"—Postnatal growth inhibition of the turbinal skeleton in brachycephalic dog breeds (*Canis lupus familiaris*). *Anat. Rec.* 304, 154–189. doi: 10.1002/ar.24422

Wilson, D. E., and Mittermeier, R. A. (2019). *Handbook of the Mammals of the World*, Vol. 9. Barcelona: Lynx Edicions.

Yohe, L. R., Hoffmann, S., and Curtis, A. A. (2018). Vomeronasal and olfactory structures in bats revealed by diceCT clarify genetic evidence of function. *Front. Neuroanat.* 12:32. doi: 10.3389/fnana.2018.00032

Zeller, U. (1987). "Morphogenesis of the mammalian skull with special reference to Tupaia," in *Morphogenesis of the Mammalian Skull (Mammalia Depicta 13)*, eds H. J. Kuhn and U. Zeller (Hamburg: Parey), 17–50.

Conflict of Interest: 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 © 2021 Ito, Tu, Eiting, Nojiri and Koyabu. 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) and the copyright owner(s) 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 Cell and Developmental Biology

METHODS
published: 13 April 2021
doi: 10.3389/fcell.2021.667296

Pescoids and Chimeras to Probe Early Evo-Devo in the Fish *Astyanax mexicanus*

Jorge Torres-Paz* and Sylvie Rétaux*

Université Paris-Saclay, CNRS, Institut des Neurosciences Paris-Saclay, Gif-sur-Yvette, France

The fish species *Astyanax mexicanus* with its sighted and blind eco-morphotypes has become an original model to challenge vertebrate developmental evolution. Recently, we demonstrated that phenotypic evolution can be impacted by early developmental events starting from the production of oocytes in the fish ovaries. *A. mexicanus* offers an amenable model to test the influence of maternal determinants on cell fate decisions during early development, yet the mechanisms by which the information contained in the eggs is translated into specific developmental programs remain obscure due to the lack of specific tools in this emergent model. Here we describe methods for the generation of pescoids from yolkless-blastoderm explants to test the influence of embryonic and extraembryonic tissues on cell fate decisions, as well as the production of chimeric embryos obtained by intermorph cell transplantations to probe cell autonomous or non-autonomous processes. We show that *Astyanax* pescoids have the potential to recapitulate the main ontogenetic events observed in intact embryos, including the internalization of mesodermal progenitors and eye development, as followed with *zic:GFP* reporter lines. In addition, intermorph cell grafts resulted in proper integration of exogenous cells into the embryonic tissues, with lineages becoming more restricted from mid-blastula to gastrula. The implementation of these approaches in *A. mexicanus* will bring new light on the cascades of events, from the maternal pre-patterning of the early embryo to the evolution of brain regionalization.

Keywords: *Zic1:GFP* lines, evo-devo, eye, gastruloids, cavefish, *Astyanax mexicanus*, ntl, chimeric embryos

OPEN ACCESS

Edited by:
Stephanie Bertrand,
UMR 7232 Biologie Intégrative des Organismes Marins (BIOM), France

Reviewed by:
Florencia Cavodeassi,
St George's, University of London, United Kingdom
Josh Gross,
University of Cincinnati, United States

****Correspondence:***
Jorge Torres-Paz
jorge.torres-paz@cnrs.fr
Sylvie Rétaux
sylvie.retaux@cnrs.fr

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *12 February 2021*
Accepted: *25 March 2021*
Published: *13 April 2021*

Citation:
Torres-Paz J and Rétaux S (2021) Pescoids and Chimeras to Probe Early Evo-Devo in the Fish Astyanax mexicanus. Front. Cell Dev. Biol. 9:667296. doi: 10.3389/fcell.2021.667296

INTRODUCTION

Emergent model organisms offer novel possibilities to unravel specific questions related to developmental and cell biology. However, some limitations, inherent to each animal system, often render difficult the implementation of novel methodologies.

Astyanax mexicanus species has thrived as an emergent model organism for evolutionary developmental biology studies (Torres-Paz et al., 2018; Jeffery, 2020). Its success in this field is due to the existence of two markedly different eco-morphotypes within the same species. *A. mexicanus* comprises river-dwelling fish populations, "surface fish," and several populations adapted to the life in caves in complete and permanent darkness, "cavefish." During cave adaptation, striking morpho-functional modifications occurred. Compared to the surface fish, the cave-adapted morphs have completely lost their eyes and pigmentation. In addition, some constructive traits have also emerged such as larger olfactory organs and more numerous facial neuromasts and taste buds, which probably contribute to a sensory compensation for the loss of the visual system (Varatharasan et al., 2009; Yoshizawa et al., 2010; Bibliowicz et al., 2013; Hinaux et al., 2016; Blin et al., 2018). Most of the morphological differences observed in the nervous system of *A. mexicanus* morphotypes have an early embryonic origin (Yamamoto and Jeffery, 2000; Pottin et al., 2011; Hinaux et al., 2016; Rétaux et al., 2016). In fact, recent evidence has shown that maternal determinants, present in the oocyte before fertilization and before zygotic developmental programs are initiated, have an important contribution to later phenotypes (Ma et al., 2018, 2020; Torres-Paz et al., 2019). Indeed, any differential composition of maternal determinants in the eggs is susceptible to lead to changes in early developmental events, such as activation of the zygotic genome, embryonic patterning and establishment of signaling centers, thus affecting later ontogenetic processes.

In fish, the extraembryonic yolk cell (of maternal origin) is an important source of inductive signals that pattern the overlying blastoderm, the embryo proper. Asymmetric segregation of maternal determinants leads to the induction of the embryonic organizer in the prospective dorsal side of the blastoderm. This symmetry breaking event will lead to localized production of different morphogens, creating gradients of signaling activity within the developing embryo. The integration of these signals by embryonic cells provides them with positional information and instructs them to follow a particular developmental program (Schier and Talbot, 2005). Thus, changes in the information contained in the oocytes, represented by maternally inherited RNAs and proteins, will affect the subsequent sequence of developmental events. Hence, *A. mexicanus* offers a unique model to test the maternal influence on embryonic development. However, methods to assess the effect of signaling centers (embryonic and extraembryonic) and the potential of cells to respond to these signals have not been developed yet in this model.

Here, we describe the implementation in *A. mexicanus* of methods used in well-established fish models to probe mechanisms of cell/tissue specification during early embryogenesis. First, we have adapted a recent method of embryonic explant culture developed in zebrafish and known as "pescoids" (i.e., gastruloids derived from fish embryonic cells) to grow the blastoderm after removal of the extraembryonic yolk cell (Fulton et al., 2020; Schauer et al., 2020). Under these conditions of altered embryonic geometry and physical constraints, the pescoids are able to recapitulate the main processes observed in intact embryos such as symmetry breaking, germ layer specification and elongation. In addition to previous reports on pescoids, here we found clear indications of mesoderm internalization and eye development. These pescoids will allow comparative analyses of gene expression in *Astyanax* morphs in the absence of yolk-derived signaling. Second, we have set up the conditions to efficiently achieve inter-morph cell transplantations at matching stages during early embryogenesis. Cell grafting have been widely performed in zebrafish embryos to test cell autonomy and potential during development, as well as to dissect lineage and timing aspects during cell specification. Grafts have also been performed between distinct species such as zebrafish and medaka to study developmental heterochronies (Fuhrmann et al., 2020). In *A. mexicanus*, inter-morphs cell transplantation will allow asking similar questions in a micro-evolutionary context. Further, the implementation of these methodologies to generate pescoids and chimeric embryos in *A. mexicanus* will help to explore the effect of embryonic and extraembryonic signals in cell decisions during early development.

MATERIALS AND METHODS

Fish and Embryo Collection

Our *A. mexicanus* colonies were obtained in 2004 from the Jeffery laboratory at the University of Maryland, College Park, United States. The surface fish stock derives from rivers in Texas, United States and the cavefish from the Pachón cave in San Luis Potosi, Mexico. Fish were since then maintained on a 12:12 hr light:dark cycle at a temperature of 22°C for cavefish and 26°C for surface fish. Reproductions were induced every other week by changes in water temperature: for cavefish temperature was increased to 26°C, and for surface fish temperature was decreased to 22°C during 3 days followed by an increase to 26°C (Elipot et al., 2014). Fish from both morphotypes spawn regularly the first and second days following the increase in temperature. Here, embryos were obtained exclusively by *in vitro* fertilization in order to ascertain synchronous early development. Embryo dechorionation was performed by enzymatic treatment with Pronase 1 mg/mL (Sigma) and embryos were maintained in Embryo Medium (EM) at 24°C. Surface and Pachón cavefish *zic1:GFP* transgenic lines used here were generated previously in the lab (Devos et al., 2019). Animals were treated according to the French and European regulations for handling of animals in research. SR's authorization for use of *Astyanax mexicanus* in research is 91–116. This work did not necessitate a protocol authorization number from the Paris Centre-Sud Ethic Committee. The animal facility of the Institute received authorization 91272105 from the Veterinary Services of Essonne, France, in 2015.

Generation of Chimeric Embryos by Cell Transplantations

Donor embryos were injected at the 1-cell stage with 3–5 nL of 1% Dextran-FITC 10,000 MW (Molecular Probes) and 0.05% Phenol Red (to see the solution) using a FemtoJet (Eppendorf). Glass pipettes for microinjection and cell transplantation were prepared on a Narishige PN-30 puller using borosilicate glass capillary (GC100F15 Harvard Apparatus LTD and B120-69-10 WPI, respectively). Microinjection pipettes were sealed at the tip and broken for opening at the moment of the injection using forceps. Cell transplantations pipettes were prepared in advance, the tip was broken at the desired internal diameter (15–30 μm) and polished using a Micropipette grinder (Narishige EG-44) at an angle of 35° in order to create a smooth needle-shaped tip. Our cell transplantation system consisted of a holder for the glass pipette (WPI) connected to a 1 mL syringe by a Teflon tubing (Narishige). Under a fluorescent dissecting microscope, labeled donor cells were aspirated into the tip of the glass pipette filled with EM, and 3–12 cells expelled into the host embryo with gentle pressure. Host embryos were let to develop in EM until fixation. In this study, isotopic and isochronic intermorphs grafts were performed (Surface animal pole cells into Cave animal pole, either at blastula or gastrula stages). Of note, targeting of cells to specific regions should be significantly more precise at gastrula stages, with the appearance of the organizer as a reference and morphological landmark for the dorsal side. *In vivo*, dextran-FITC labeled cells in grafts could be visualized at least up to 1 dpf (not tested later in this study).

Generation of Pescoids

Astyanax mexicanus pescoids were produced following a recent description in zebrafish (Fulton et al., 2020; **Figure 1A**). Briefly, at the 512-1K cell stage the yolk was carefully removed from embryos using eyebrows knives. Blastoderm explants were cultured until the corresponding 11 h post-fertilization (hpf) at 24°C in L15 medium (Gibco) supplemented with 3% Fetal Bovine Serum (FBS, Biosera). For cultures maintained for longer than 7 h Penicillin-Streptomycin were added to the culture media (1x dilution, P4333, Sigma).

Histology and Imaging

Colorimetric ISH was performed as previously described (Pottin et al., 2011). Digoxigenin-labeled riboprobe was prepared using PCR products as templates. *no-tail* (*ntl*) cDNA was obtained from our ESTs library (accession number ARA0ABA99YL22). Procedure for revelation of FITC-labeled donor cells combined with fluorescent ISH was adapted from our ISH protocol (Alié et al., 2018). Following the hybridization with the Digoxigenin-labeled ISH probe, embryos were first processed for the revelation of the Dextran-FITC grafted cells: they were incubated for 1 h in blocking solution (Tris 0.1M pH 7.5, NaCl 150 mM, Tween-20 0.1 and 5% blocking reagent Roche) and then with POD-conjugated anti-FITC antibody (11426346910; Roche, 1/400) diluted in blocking solution overnight. Embryos were washed in PBS/Tween 0.1% (PBST) 10 times for 10 min each time and incubated for 30 min at room temperature with TAMRA-tyramide 1/1000. Peroxidase activity was activated by H_2O_2 (0.003%, Sigma) for 1 h and samples were washed again 10 times for 10 min in PBST. Revelation of the dig-labeled ISH probe was then performed using an anti-Digoxigenin antibody coupled to POD (11207733910; Roche, 1/400) and revealed using a FITC-tyramide (1/400). After several washes in PBST embryos were stained with DAPI (10236276001, Sigma) at a final concentration of 1 mg/ml, overnight at 4°C, and washed in PBS before dissection and mounting (Vectashield, Vector Laboratories).

Immunohistochemistry was performed as previously described (Blin et al., 2018) using a primary anti-GFP antibody at a dilution 1/500 (GFP-1020, Aves Labs) and a secondary Alexa Fluor 488-coupled antibody at a dilution 1/500 (Anti-chicken, A-11039, Invitrogen). Fluorescently labeled embryos were counterstained with DAPI. Embryos stained by colorimetric ISH were imaged on a Nikon AZ100 multizoom macroscope. Confocal acquisitions were done on a Leica-SP8 confocal microscope using the Leica Application Suite software. Images processing and quantifications were done on Fiji and statistical tests (Mann-Whitney non-parametric tests) were performed on GraphPad Prism.

RESULTS AND DISCUSSION

Generation of Astyanax Pescoids

Recent advances in the field of gastruloids have highlighted the robustness of animal development and the key steps taking place during this process. In zebrafish pescoids (yolk-less blastoderm explants) the main aspects of early development observed in intact embryos are recapitulated. Symmetry breaking, axis elongation and neural specification occur despite the absence of extraembryonic signaling (Fulton et al., 2020; Schauer et al., 2020). In fish, these events are controlled maternally (Marlow, 2020; Solnica-Krezel, 2020) and *A. mexicanus* with its two morphotypes has become an excellent model to challenge the role of maternal determinants in embryogenesis (Ma et al., 2018, 2020; Torres-Paz et al., 2019).

Pescoids derived from embryos of both *Astyanax* surface and cavefish embryos (surfoids and caveoids, respectively) developed similarly to those described in zebrafish. After removal of the vitellus at the 256-1K cell stage (**Figure 1B**), blastoderm explants sealed the wound and became rounded during the first hour post culture (hpc, equivalent to 5–6 hpf, **Figure 1C**). Then, during the next 3–4 h in culture a cavity was formed (8–9 hpf, **Figure 1D**), which may correspond to a "blastocoel" (Schauer et al., 2020). After 5 hpc (9 hpf), the first signs of axis elongation were observed in surfoids and caveoids (**Figures 1E–G**). The shape of the elongated pescoids was asymmetrical and pear-shaped, with a narrow tip at one end and a larger rounded form at the opposite extremity. The extent of elongation was quantified by calculating the aspect ratio (length/width, **Figure 1F**) at three timepoints up to the end of the culture at stages corresponding to tailbud/early somitogenesis (9, 10, and 11 hpf, **Figures 1F–H**). Both surfoids and caveoids underwent significant elongation during this time window. At the three stages analyzed, however, we found reduced elongation of caveoids compared to surfoids (**Figure 1G**). Moreover, we observed inter clutch variability in the extent of elongation in pescoids derived from different cavefish females, but much less so between different surface fish females, and always with higher mean length values in surface-derived pescoids as shown at 11 hpf (**Figure 1H**). These results are counterintuitive considering our previous observations of more precocious axial mesoderm extension in cavefish compared to surface embryos during gastrulation (Torres-Paz et al., 2019). Reduced elongation in cavefish pescoids may originate from deficiencies in the regulative properties of embryonic cells and/or the maternal pre-patterning at the blastula stages (Fulton et al., 2020). Insights on maternal control of early patterning and morphogenesis will come from pescoids obtained from F1 reciprocal hybrid embryos (Ma et al., 2018; Torres-Paz et al., 2019).

To get further insights into the elongation process, we studied the expression of the mesodermal marker *ntl* in *Astyanax* pescoids and found a pattern strikingly reminiscent

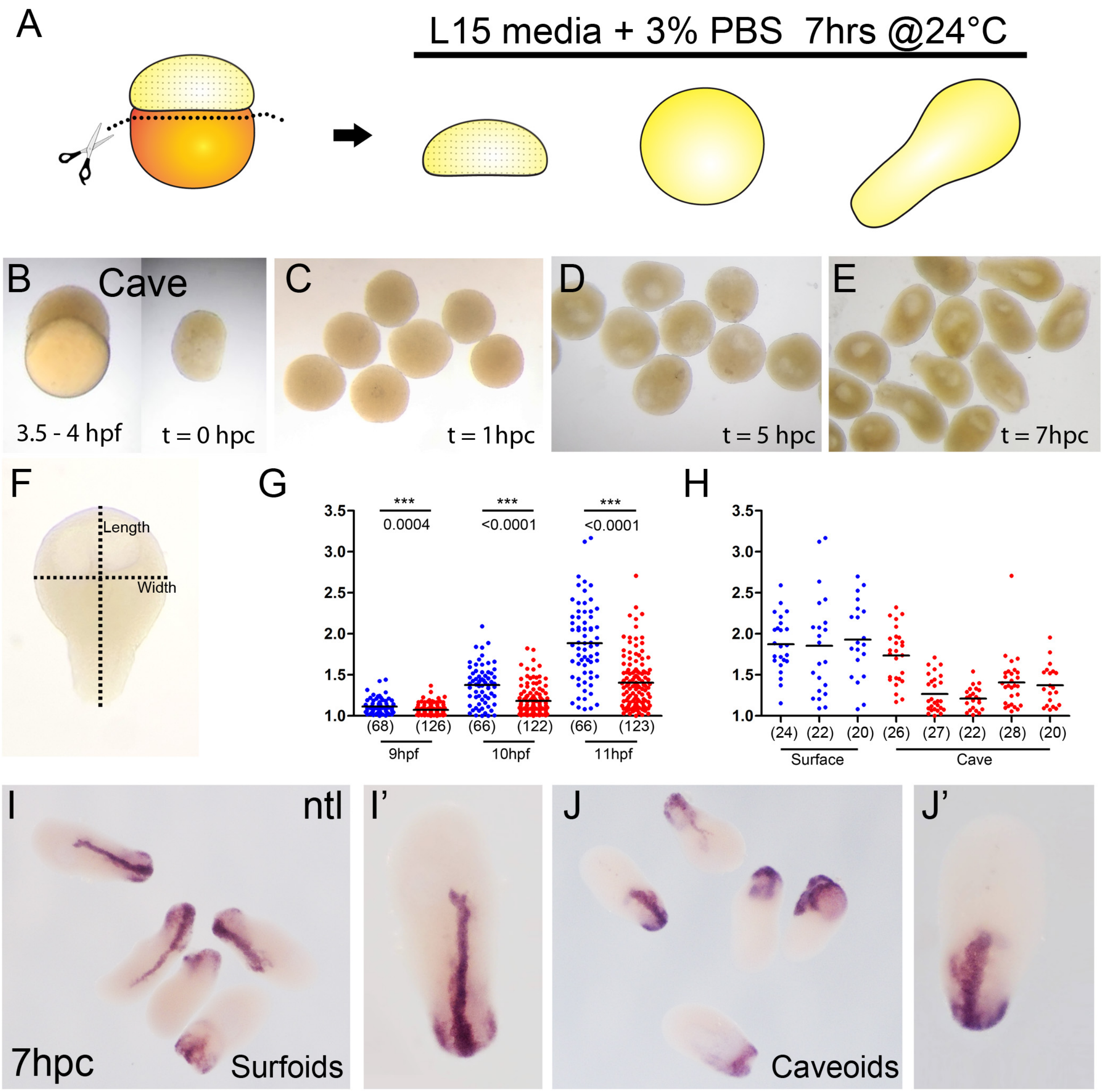


FIGURE 1 | Generation of pescoids in *A. mexicanus*. Procedure for the generation of *Astyanax* pescoids **(A)**. Intact cavefish embryo at 1K-cell stage [**(B), left**] and after yolk removal [**(B), right**]. Development of caveoids after 1, 5, and 7 hpc [**(C–E)**, respectively]. Measurement of the length and width on a pescoid to calculate the aspect ratio value as length/width **(F)**. Measurement of aspect ratio of surface (blue) and cave (red) pescoids at 9, 10, and 11 hpf **(G)**. Statistical differences (***) between the two morphs are indicated with the corresponding p values obtained by Mann-Whitney tests at every time point and the numbers in parentheses indicate the number of pescoids analyzed **(G)**. Aspect ratio measurement at 11 hpf in pescoids derived from individual females shows variability in cavefish (3 different females for surfoids, 5 different females for caveoids) **(H)**. Black lines indicate the mean value **(G,H)**. Expression of *ntl* in surfoids **(I,I')** and caveoids **(J,J')** after 7 hpc.

of the developing notochord observed in intact fish embryos (**Figures 1I,J**). Microscopic observations suggested that cells expressing *ntl* had been internalized, an aspect of fish gastruloids that may have been overlooked. We thus compared the expression of *ntl* in confocal acquisitions after fluorescent ISH in intact embryos at tailbud stage and in pescoids at equivalent stages (**Figure 2**). In confocal reconstructed sections of pescoids and control embryos, we observed similar organization of *ntl* expressing cells, always underneath an overlying layer of superficial cells, that we interpreted as ectoderm cells (**Figure 2**, arrows). Thus, these data confirm a conserved internalization process of axial mesoderm in explants despite the absence of vitellus. These observations highlight the robustness of cellular processes during vertebrate gastrulation.

Studies in zebrafish have shown that asymmetric translocation of maternal determinants in the yolk leading to the dorsal

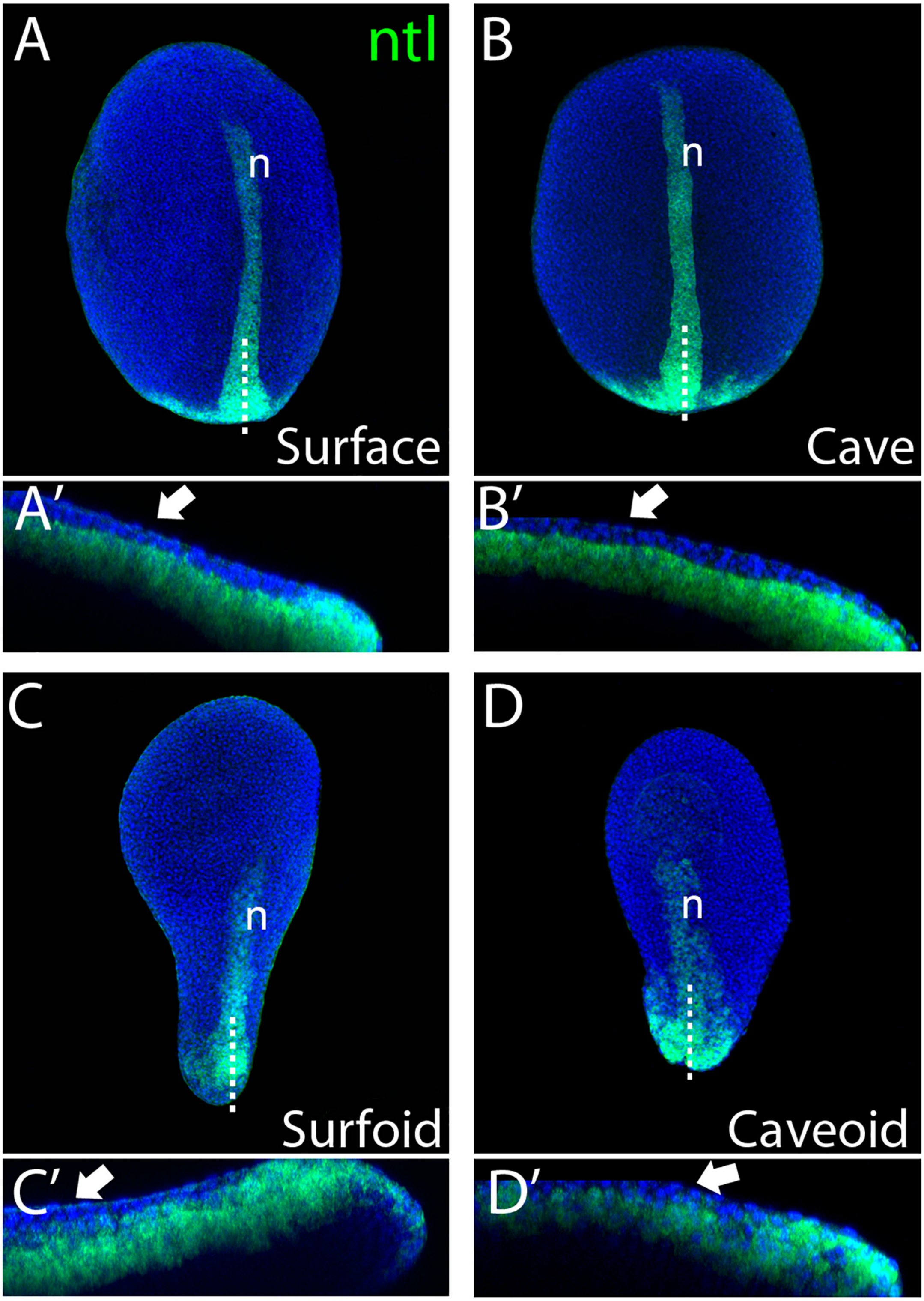


FIGURE 2 | Internalization of mesoderm in *Astyanax* pescoids. Expression of *ntl* in intact surface and cavefish embryos at 10 hpf [**(A,B)**, respectively], and in surfoid and caveoid after 7 hpc [**(C,D)**, respectively]. Embryos are oriented in dorsal views, anterior to the top. For each panel, bottom images **(A'–D')** are confocal reconstructions at the indicated level (dotted lines), anterior to the left and dorsal to the top. The white arrows indicate ectodermal cells overlying *ntl*-expressing mesodermal cells and n indicates the notochord.

determination occur as early as the 16–32 cell stage (Jesuthasan and Strähle, 1997). The activation of the zygotic genome, a process dependent on the maternal transcriptomic machinery, starts around the 64 cell stage (Chan et al., 2019), whereas clear zygotic transcription is observed from the 512 cell stage. Concurrently, at this stage, called the mid-blastula transition, embryonic cell cycles become asynchronous and the extra-embryonic yolk syncytial layer (YSL) is formed at the interphase between the yolk and the blastoderm (Kimmel et al., 1995). Similarly to zebrafish, in *Astyanax* the mid-blastula transition takes place around the 512-1K cell stage (Hinaux et al., 2011). Here, the explant experiments were performed between the 128-1K cell stage, i.e., the end of the maternal-to-zygotic transition, thus it must be taken into account that maternally-derived pre-patterning already exists in the cultured blastoderms. Consequently, in order to gain insights into the temporal sequence of maternal patterning events, the generation of even earlier explants must be considered. The comparative analysis of earlier explants in the two *Astyanax* morphs will allow dissecting precisely the timing and the impact of the maternal contributions to developmental evolution. Recently, it was shown in zebrafish that the extra-embryonic YSL layer does not form properly in yolkless-explants, yet correctly developing pescoids can be obtained even from very precocious 64 cell stage embryos (Schauer et al., 2020). This indicates that pescoids can develop into embryo-like structures in the absence YSL-derived signals (Rodaway et al., 1999; Chen and Kimelman, 2000). Conversely, animal caps explants are able to develop into structures similar to pescoids only if Nodal and downstream planar cell polarity signaling pathways are active (Williams and Solnica-Krezel, 2020), indicating that Nodal activity in pescoids must come from marginal cells. Nodal signaling is necessary for the induction of endomesodermal fates at the blastoderm margin (Schier et al., 1997; Vopalensky et al., 2018). Visual inspection of the mesodermal marker *ntl* expression in our pescoids at the different states of elongation clearly showed that the point where *ntl*-expressing cells are internalized corresponded to the marginal zone, where the wound closed (not shown). Thus, the wounded margin in the blastoderm explants would be topologically and functionally equivalent to the blastopore in intact embryos, being both the source of Nodal signaling and the point where endomesoderm is internalized. Given the previously described differences in organizer formation and mesoderm internalization between the two *Astyanax* morphs embryos (Torres-Paz et al., 2019), pescoids will prove powerful tools to study the origin and outcomes of these processes.

Eye Development in *Astyanax* Pescoids

Our observations of axial elongation and the overtly normal formation of the notochord in pescoids (**Figure 2**) made us wonder if eye development could also occur in these yolkless explants. We recently generated *zic1:GFP* transgenic knock-in reporter lines in both *Astyanax* morphotypes in order to visualize eye morphogenesis (Devos et al., 2019). We took advantage of these transgenic fish to evaluate directly eye development in pescoids. Blastoderm explants from *zic1:GFP* embryos were let to develop *in vitro* until the equivalent of 24 hpf, i.e., when the optic cups are formed in surface and cave embryos (**Figures 3A,B**). Despite some "body axes" malformations in these explants (not shown), we observed GFP reporter expression in discrete regions within the prospective pescoid head (**Figures 3C–E**). Strikingly, we were able to identify distinct GFP expression domains that we interpreted as corresponding to the optic (arrows, **Figure 3**) and anterior telencephalic ("t," **Figure 3**) *zic1*-expressing tissues, respectively, like in intact embryos (**Supplementary Figure 1**). We cannot rule out that, in some pescoids, the separation of the telencephalic and optic GFP expressing cells could be perturbed and result in a single GFP positive domain containing intermingled cells with mixed identities (**Figure 3D**). Further studies analyzing cell movements during pescoids "neurulation" will be needed to address this question.

In embryos, the size of the optic cup is known to be smaller in cavefish (Yamamoto and Jeffery, 2000; Alunni et al., 2007). We observed the same tendency in the *zic1:GFP* transgenics (**Figures 3A,B**, Surface = 23,557 ± 1,336 μm^2, *n* = 4, mean ± standard deviation, n; Cave = 15,169 ± 2,395 μm^2, *n* = 3; −35% in cavefish embryos; see also Devos et al., 2019). In pescoids, the extension of the optic domain delimited by reporter expression varied importantly, when present and identifiable, in the two morphotypes, but it was smaller in caveoids (**Figures 3C–E**, Surfoids = 15,998 ± 6,329 μm^2, *n* = 3; Caveoids = 11,355 ± 4,460 μm^2, *n* = 4; −30% in caveoids). Thus, on the small sample analyzed the difference in size of the optic tissue mirrors the situation in whole embryos, suggesting a conservation of the control of organ size in pescoids. Surprisingly, in two surfoid specimens we even found a high degree of retinal morphogenesis (**Figures 3E,E1–E4**), with clear indications of epithelial folding (arrow, **Figure 3E**) and neural tissue contacting adjacent non-neural ectoderm (arrowheads, **Figures 3E1–E4**). These observations further highlight the robustness of developmental processes occurring in gastruloids.

Eye development starts with the specification of eye precursor cells in the anterior neural plate (Zuber et al., 2003; Cavodeassi et al., 2005), followed by complex morphogenesis processes that involve coordinated movements within the neural plate (Cavodeassi et al., 2005; Rembold et al., 2006; Picker et al., 2009; Kwan et al., 2012; Ivanovitch et al., 2013) that are instructed by midline signaling (Macdonald et al., 1995; García-Calero et al., 2008; Gordon et al., 2018). Our observations of eye-like embryonic tissue in *Astyanax* pescoids demonstrate the potential of gastruloid systems to engage into complex morphogenesis. The wide spectrum of optic phenotypes observed in our pescoids will allow a better understanding of minimal requirements for particular developmental processes and the interdependency of different embryonic tissues during morphogenesis.

Generation of Chimeric Embryos by Cell Grafting

Cell transplantation methodologies have been widely used in zebrafish to test cell-autonomy during embryogenesis in different experimental contexts (Fauny et al., 2009; Cavodeassi et al., 2013; Giger et al., 2016). *A. mexicanus* with its two eco-morphotypes offers a great opportunity to test autonomy during

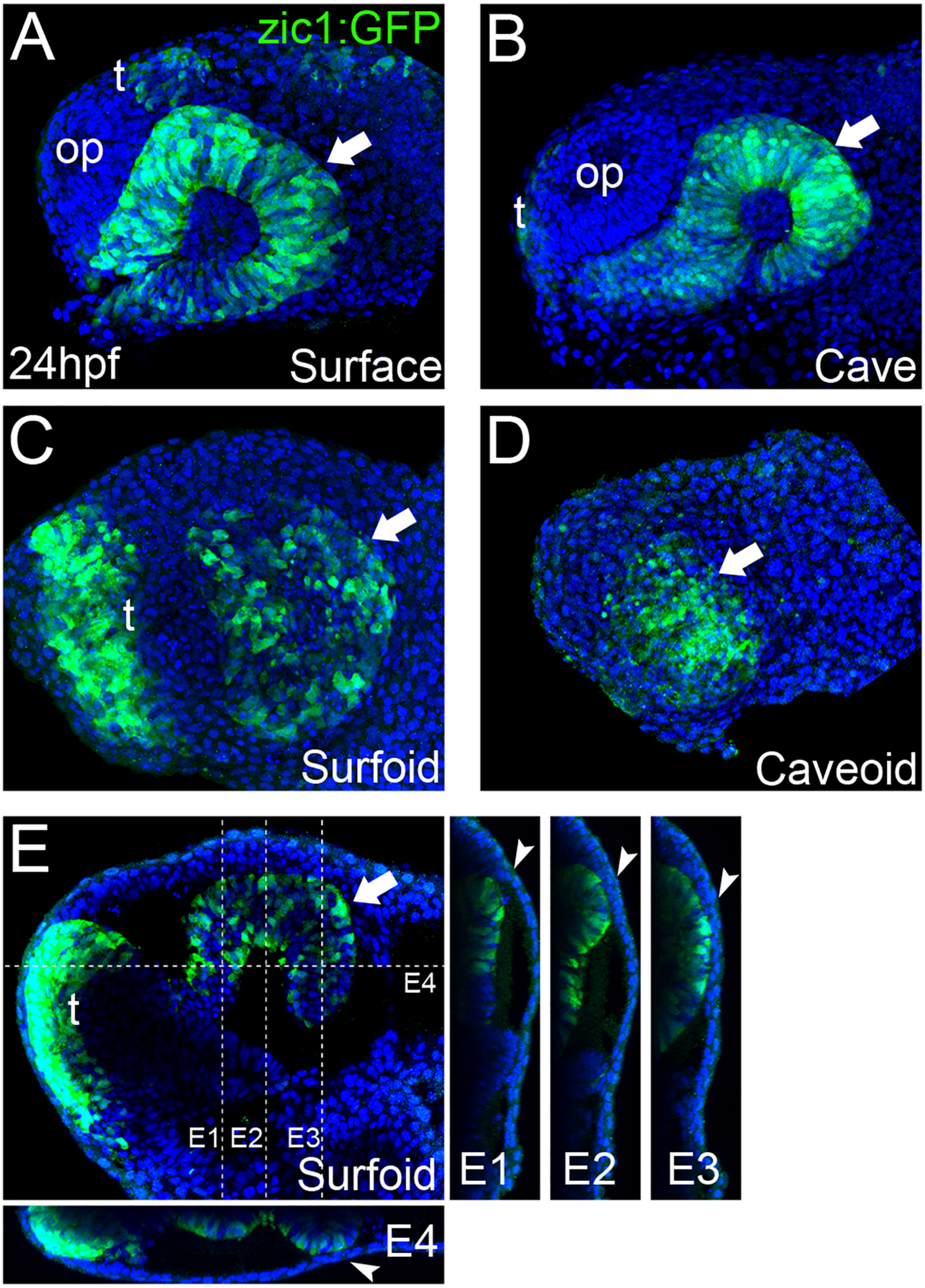

A
zic1:GFP
t
op
24hpf
Surface
B
op
t
Cave
C
t
Surfoid
D
Caveoid
E
t
E4
E1 E2 E3
Surfoid
E1
E2
E3
E4

FIGURE 3 | Eye development in *Astyanax* pescoids. GFP-reporter expression in *zic1:GFP* transgenic surface and cavefish intact embryos [**(A,B)**, respectively], and surfoid **(C,E)** and caveoid **(D)** at 24 hpf. Embryos and pescoids are oriented in lateral views, anterior to the left. Arrows indicate the GFP-expressing optic tissue; op, olfactory placode; t, telencephalon. Confocal reconstructions at the indicated levels in **(E)** are shown **(E1–E4)**. Arrowheads **(E1–E4)** indicate the contact zone between neural and non-neural ectoderm.

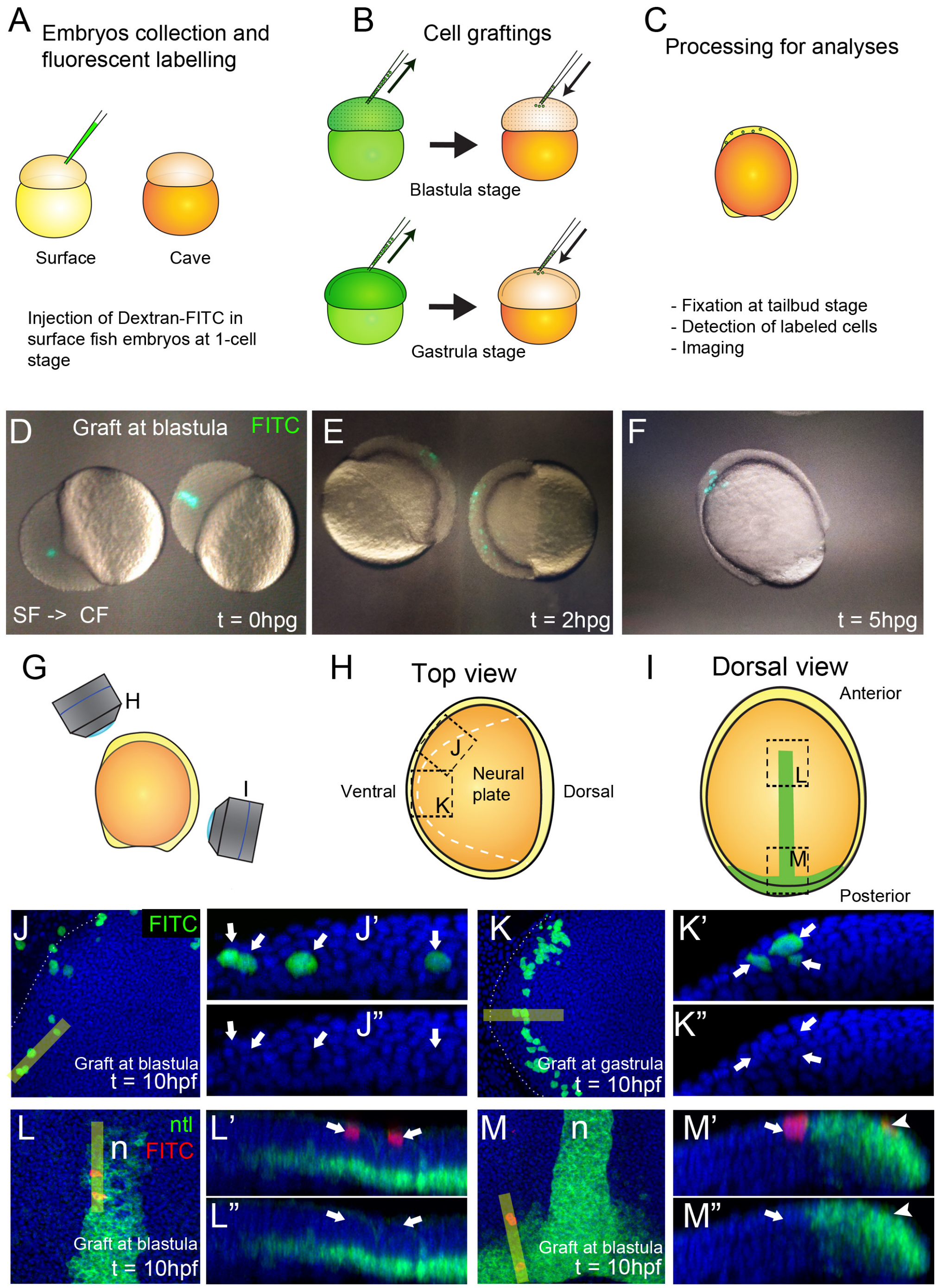

A
Embryos collection and fluorescent labelling
Surface
Cave
Injection of Dextran-FITC in surface fish embryos at 1-cell stage
B
Cell graftings
Blastula stage
Gastrula stage
C
Processing for analyses
- Fixation at tailbud stage
- Detection of labeled cells
- Imaging
D
Graft at blastula
FITC
SF -> CF
t = 0hpg
E
t = 2hpg
F
t = 5hpg
G
H
I
H
Top view
Ventral
J
K
Neural plate
Dorsal
I
Dorsal view
Anterior
L
M
Posterior
J
FITC
Graft at blastula
t = 10hpf
J'
J"
K
Graft at gastrula
t = 10hpf
K'
K"
L
n
ntl
FITC
Graft at blastula
t = 10hpf
L'
L"
M
n
Graft at gastrula
Graft at blastula
t = 10hpf
M'
M"

FIGURE 4 | Generation of chimeric embryos in *A. mexicanus* by inter-morph cell transplantation. Procedure for cell grafting between surface and cavefish embryos **(A–C)**. Live cavefish (host) embryos that have been transplanted with FITC-labeled surface cells (green) at the blastula stage, photographed immediately after the grafting [0 hpg, **(D)**], or after 2 hpg **(E)** and 5 hpg **(F)**. Fluorescent labeled embryos at tailbud stage were imaged in a confocal microscope in a view from the top **(G,H)** and from the dorsal side **(G,I)**. Black dashed lines indicate the orientation of the images shown in panels **(J–M)**. Dashed white line indicate the neural plate border **(H,J,K)**. Confocal images of dissected cavefish embryos at 10 hpf at the level of the anterior neural plate **(J, K)**. Labeling in green corresponds to surface fish cells grafted at blastula **(J)** and gastrula **(K)** stages. Respective reconstructed sections were obtained at the levels indicated in yellow bars and grafted cells are indicated with arrows **(J′,J″,K′,K″)**. Confocal images of dissected cavefish embryos at 10 hpf stained for *ntl* [green, **(I,L,M)**], at the level of the anterior notochord **(L)** and the tail bud **(M)**. Labeling in red corresponds to surface fish cells grafted at blastula stage and labeling in greencorresponds to *ntl* expression in the notochord. Respective reconstructed sections were obtained at the levels indicated in yellow bars **(L′,L″,M′,M″)**. Arrows **(L′,L″,M′,M″)** indicate superficial ectodermal grafted cells. Arrowheads **(M′,M″)** indicate a cell in the mesodermal domain expressing *ntl*. Images in panels **(J,K)** are top views, anterior to the left. Panels **(L,M)** are dorsal views, anterior to the top. Reconstructions are oriented with anterior to the left and dorsal on top.

cell fate specification through intermorph transplantations. Hence, elegant transplantation experiments of the lens from one morph into the optic cup of the other morph at larval stages have revealed the autonomy of the cavefish lens apoptotic process and its triggering role in cavefish eye degeneration (Yamamoto and Jeffery, 2000). Neural crest cell transplantations have been performed as well at 2 dpf to address their role in cavefish pigmentation defect (Yoshizawa et al., 2018). However, the *Astyanax* model has not succeeded yet for early cell transplantations aimed at studying precocious embryogenesis, mainly due to technical challenges that must be circumvented. A major challenge is the simultaneous collection of embryos of both morphotypes at equivalent developmental stage. *A. mexicanus* reproduce in the dark (Simon et al., 2019), thus in order to obtain early developing embryos to work with during the day, we inverted the circadian cycle of fish in a special fish room dedicated to reproduction in our facility. In addition, to obtain early embryos developing synchronously, *in vitro* fertilizations must be performed using ready-to-spawn females. Hence, mating behavior was monitored during the days following the induction of reproductions with a camera in the circadian-inverted fish room illuminated with a dim red-light. Inductions of surface and cavefish were then coordinated in order to find spawning females of the two morphs at the same time.

In this work surface fish embryos were labeled with dextran-FITC at the one-cell stage (donors), and cells were grafted isochronically into unlabeled cavefish embryos (hosts) at two developmental time points, the mid-blastula transition (512-2K cell stage) and the onset of gastrulation (30–50% epiboly) (**Figures 4A,B**). As a source of donor cells we choose the embryonic animal pole, in order to compare to fate maps studies performed in zebrafish and showing that ectodermal precursors (including neurectoderm) arise from this field, whereas endomesodermal precursors derive from more marginal cells (Woo and Fraser, 1997; Schier and Talbot, 2005). After transplantation of fluorescently labeled cells (**Figure 4D**), embryonic development was observed to occur normally during the following hours post grafting (hpg), with labeled cells integrated in the embryo (**Figures 4E,F**). After fixation and methanol storage of chimeric embryos, FITC fluorescence in labeled grafted cells was completely lost (not shown), rendering necessary a revelation through immunohistochemistry (**Figure 4C**). Fluorescent revelation of labeled cells with FITC-coupled tyramide was avoided because an extensive bleed through of fluorescence to channels at lower wavelengths was observed (not shown). Instead, revelation with TAMRA-coupled tyramide, whose fluorescent excitation/emission occurs at higher wavelengths (557 and 583 nm, respectively) than FITC (495 and 521 nm, respectively), did not produce bleed through to lower wavelengths channels (**Figures 4G′,G′′,H′,H′′**).

We compared the organization and repartition of grafted cells in host embryos at the two transplantation stages, blastula versus gastrula, and observed clear differences between the two conditions. Clones transplanted at blastula stage were distributed extensively throughout the embryos and in a disorganized and scattered manner. On the other hand, cells transplanted at gastrula stages were restricted in space and in some cases showed clear signs of symmetry (compare **Figures 4E,F**). Using the positional information and the expression of *ntl* in the notochord as reference (**Figure 4I**), we observed that cells grafted at blastula stages were able to produce both ectodermal (n = 15/15 embryos, superficial cells in **Figures 4L,L′,L′′**) and endomesodermal (n = 6/15 embryos, *ntl* positive cell in **Figures 4M,M′,M′′**) derivatives in chimeric embryos. On the other hand, grafts performed at gastrula stages gave rise to only ectodermal cells (n = 24/24 embryos, **Figures 4J,J′,J′′,4K,K′,K′′**). These data were consistent with an expected progressive lineage restriction from mid-blastula to early gastrulation stages. Confocal reconstructions suggested a correct integration of transferred cells in the developing host tissues. Similar results were also observed in reciprocal experiments, i.e., transplants of cavefish donor cells into a surface fish host (not shown). We also found that surface fish embryonic cells were able to differentiate into pigmented cells in a cavefish host (**Supplementary Figure 2**) confirming the correct integration and differentiation of donor cells in chimeric embryos and illustrating a typical cell-autonomous process.

Intermorph grafting will shed light on the cell autonomy and the effect of the embryonic signaling environment on previously described heterochronies, heterotopies and differences of gene expression levels during development of *Astyanax* morphs (Yamamoto et al., 2004; Pottin et al., 2011; Hinaux et al., 2016; Torres-Paz et al., 2019). The combination of these grafting methods with the use of transgenic reporter lines such as the cavefish and surface fish *zic1:GFP* lines (Devos et al., 2019), will allow the detailed investigation of intrinsic and extrinsic factors implicated in eye specification and degeneration.

CONCLUSION

Implementation and optimization of new methods in emergent model systems is fundamental for tackling novel scientific questions. Here we describe the methodology and potential applications of cellular techniques to generate yolk-free pescoids and chimeric embryos in *Astyanax mexicanus*. These methods will allow the characterization of developmental states during cell lineages differentiation in embryogenesis. In addition, these techniques will push forward genomic and cellular approaches to understand the key steps during eye development and degeneration in cavefish.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The animal study was reviewed by the French and European regulations for handling of animals in research. SR's authorization for use of Astyanax mexicanus in research is 91-116. This work did not necessitate a protocol authorization number from the Paris Centre-Sud Ethic Committee. The animal facility of the Institute received authorization 91272105 from the Veterinary Services of Essonne, France, in 2015.

AUTHOR CONTRIBUTIONS

JT-P performedexperiments. JT-P and SR conceived experiments, analyzed data, and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

Grant support was received from UNADEV/AVIESAN/Vision and Retina France, CNRS and Becas Chile.

ACKNOWLEDGMENTS

We thank Stéphane Père, Victor Simon, and Krystel Saroul for taking care of our *Astyanax* colony, Maryline Blin for technical support for ISH protocols and François Agnès and all other current and past members of the DECA team for fruitful discussions and important suggestions. We also thank Ben Steventon for his advices on the protocol for generation of pescoids, and Nicolas David and Lazaro Centanin for technical advices on cell transplantation methods.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.667296/full#supplementary-material

REFERENCES

Alié, A., Devos, L., Torres-Paz, J., Prunier, L., Boulet, F., Blin, M., et al. (2018). Developmental evolution of the forebrain in cavefish, from natural variations in neuropeptides to behavior. *ELife* 7:e32808. doi: 10.7554/eLife.32808

Alunni, A., Menuet, A., Candal, E. V. A., Jeffery, W. R., and Re, S. (2007). Developmental mechanisms for retinal degeneration in the blind cavefish *Astyanax mexicanus. J. Comp. Neurol.* 505, 221–233. doi: 10.1002/cne

Bibliowicz, J., Alié, A., Espinasa, L., Yoshizawa, M., Blin, M., Hinaux, H., et al. (2013). Differences in chemosensory response between eyed and eyeless *Astyanax mexicanus* of the Rio Subterráneo cave. *EvoDevo* 4:25. doi: 10.1186/2041-9139-4-25

Blin, M., Tine, E., Meister, L., Elipot, Y., Bibliowicz, J., Espinasa, L., et al. (2018). Developmental evolution and developmental plasticity of the olfactory epithelium and olfactory skills in Mexican cavefish. *Dev. Biol.* 441, 242–251. doi: 10.1016/j.ydbio.2018.04.019

Cavodeassi, F., Carreira-Barbosa, F., Young, R. M., Concha, M. L., Allende, M. L., Houart, C., et al. (2005). Early stages of zebrafish eye formation require the coordinated activity of Wnt11, Fz5, and the Wnt/β-catenin pathway. *Neuron* 47, 43–56. doi: 10.1016/j.neuron.2005.05.026

Cavodeassi, F., Ivanovitch, K., and Wilson, S. W. (2013). Eph/Ephrin signalling maintains eye field segregation from adjacent neural plate territories during forebrain morphogenesis. *Development* 140, 4193–4202. doi: 10.1242/dev.097048

Chan, S. H., Tang, Y., Miao, L., Darwich-Codore, H., Vejnar, C. E., Beaudoin, J. D., et al. (2019). Brd4 and P300 confer transcriptional competency during zygotic genome activation. *Dev. Cell* 49, 867–881.e8. doi: 10.1016/j.devcel.2019.05.037

Chen, S. R., and Kimelman, D. (2000). The role of the yolk syncytial layer in germ layer patterning in zebrafish. *Development* 127, 4681–4689.

Devos, L., Klee, F., Edouard, J., Simon, V., Legendre, L., Khallouki, N. E., et al. (2019). Morphogenetic and patterning defects explain the coloboma phenotype of the eye in the Mexican cavefish. *[Preprint] BioRxiv* doi: 10.1101/698035

Elipot, Y., Legendre, L., Père, S., Sohm, F., and Rétaux, S. (2014). *Astyanax* transgenesis and husbandry: how cavefish enters the laboratory. *Zebrafish* 11, 291–299. doi: 10.1089/zeb.2014.1005

Fauny, J. D., Thisse, B., and Thisse, C. (2009). The entire zebrafish blastula-gastrula margin acts as an organizer dependent on the ratio of nodal to BMP activity. *Development* 136, 3811–3819. doi: 10.1242/dev.039693

Fuhrmann, J. F., Buono, L., Adelmann, L., Martinez-Morales, J. R., and Centanin, L. (2020). Genetic developmental timing revealed by inter-species transplantations in fish. *Development* 147:dev192500. doi: 10.1242/dev.192500

Fulton, T., Trivedi, V., Attardi, A., Anlas, K., Dingare, C., Arias, A. M., et al. (2020). Axis specification in zebrafish is robust to cell mixing and reveals a regulation of pattern formation by morphogenesis. *Curr. Biol.* 30, 2984–2994.e3. doi: 10.1016/j.cub.2020.05.048

García-Calero, E., Fernández-Garre, P., Martínez, S., and Puelles, L. (2008). Early mammillary pouch specification in the course of prechordal ventralization of the forebrain tegmentum. *Dev. Biol.* 320, 366–377. doi: 10.1016/j.ydbio.2008.05.545

Giger, F. A., Dumortier, J. G., and David, N. B. (2016). Analyzing in vivo cell migration using cell transplantations and time-lapse imaging in zebrafish embryos. *J. Vis. Exp.* 2016:e53792. doi: 10.3791/53792

Gordon, H. B., Lusk, S., Carney, K. R., Wirick, E. O., Murray, B. F., and Kwan, K. M. (2018). Hedgehog signaling regulates cell motility and optic fissure and stalk formation during vertebrate eye morphogenesis. *Development* 145:dev165068. doi: 10.1242/dev.165068

Hinaux, H., Devos, L., Blin, M., Elipot, Y., Bibliowicz, J., Alié, A., et al. (2016). Sensory evolution in blind cavefish is driven by early embryonic events during gastrulation and neurulation. *Development* 143, 4521–4532. doi: 10.1242/dev.141291

Hinaux, H., Pottin, K., Chalhoub, H., Père, S., Elipot, Y., Legendre, L., et al. (2011). A developmental staging table for *Astyanax mexicanus* surface fish and pachón cavefish. *Zebrafish* 8, 155–165. doi: 10.1089/zeb.2011.0713

Ivanovitch, K., Cavodeassi, F., and Wilson, S. W. (2013). Precocious acquisition of neuroepithelial character in the eye field underlies the onset of eye morphogenesis. *Dev. Cell* 27, 293–305. doi: 10.1016/j.devcel.2013.09.023

Jeffery, W. R. (2020). *Astyanax* surface and cave fish morphs. *EvoDevo BioMed. Cent.* 11:14. doi: 10.1186/s13227-020-00159-6

Jesuthasan, S., and Strähle, U. (1997). Dynamic microtubules and specification of the zebrafish embryonic axis. *Curr. Biol.* 7, 31–42. doi: 10.1016/S0960-9822(06)00025-X

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310. doi: 10.1002/aja.1002030302

Kwan, K. M., Otsuna, H., Kidokoro, H., Carney, K. R., Saijoh, Y., and Chien, C. B. (2012). A complex choreography of cell movements shapes the vertebrate eye. *Development* 139, 359–372. doi: 10.1242/dev.071407

Ma, L., Ng, M., Shi, J., Gore, A. V., Castranova, D., Weinstein, B. M., et al. (2020). Evolutionary changes in left-right visceral asymmetry in *Astyanax* cavefish. *[Preprint] BioRxiv* doi: 10.1101/2020.05.15.098483

Ma, L., Strickler, A. G., Parkhurst, A., Yoshizawa, M., Shi, J., and Jeffery, W. R. (2018). Maternal genetic effects in *Astyanax* cavefish development. *Dev. Biol.* 441, 209–220. doi: 10.1016/J.YDBIO.2018.07.014

Macdonald, R., Barth, K. A., Xu, Q., Holder, N., Mikkola, I., and Wilson, S. W. (1995). Midline signalling is required for pax gene regulation and patterning of the eyes. *Development* 121, 3267–3278.

Marlow, F. L. (2020). Setting up for gastrulation in zebrafish. *Curr. Top. Dev. Biol.* 136, 33–83. doi: 10.1016/bs.ctdb.2019.08.002

Picker, A., Cavodeassi, F., Machate, A., Bernauer, S., Hans, S., Abe, G., et al. (2009). Dynamic coupling of pattern formation and morphogenesis in the developing vertebrate retina. *PLoS Biol.* 7:e1000214. doi: 10.1371/journal.pbio.1000214

Pottin, K., Hinaux, H., and Rétaux, S. (2011). Restoring eye size in *Astyanax mexicanus* blind cavefish embryos through modulation of the Shh and Fgf8 forebrain organising centres. *Development* 2476, 2467–2476. doi: 10.1242/dev.054106

Rembold, M., Loosli, F., Adams, R. J., and Wittbrodt, J. (2006). Individual cell migration serves as the driving force for optic vesicle evagination. *Science* 313, 1130–1134. doi: 10.1126/science.1127144

Rétaux, S., Alié, A., Blin, M., Devos, L., Elipot, Y., and Hinaux, H. (2016). "neural development and evolution in *astyanax mexicanus*: comparing cavefish and surface fish brains," in *Biology and Evolution of the Mexican Cavefish*, eds A. C. Keen, M. Yoshizawa, and S. E. McGaugh (Amsterdam: Elsevier Inc), 227–244. doi: 10.1016/B978-0-12-802148-4.00012-8

Rodaway, A., Takeda, H., Koshida, S., Broadbent, J., Price, B., Smith, J. C., et al. (1999). Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-β family signals and discrimination of mesoderm and endoderm by FGF. *Development* 126, 3067–3078.

Schauer, A., Pinheiro, D., Hauschild, R., and Heisenberg, C. P. (2020). Zebrafish embryonic explants undergo genetically encoded self-assembly. *ELife* 9:e55190. doi: 10.7554/elife.55190

Schier, A. F., Neuhauss, S. C., Helde, K. A., Talbot, W. S., and Driever, W. (1997). The one-eyed pinhead gene functions in mesoderm and endoderm formation in zebrafish and interacts with no tail. *Development* 124, 327–342.

Schier, A. F., and Talbot, W. S. (2005). Molecular genetics of axis formation in zebrafish. *Ann. Rev. Genet*. 39, 561–613. doi: 10.1146/annurev.genet.37.110801.143752

Simon, V., Hyacinthe, C., and Rétaux, S. (2019). Breeding behavior in the blind Mexican cavefish and its river-dwelling conspecific. *PLoS One* 14:e0212591. doi: 10.1371/journal.pone.0212591

Solnica-Krezel, L. (2020). Maternal contributions to gastrulation in zebrafish. *Curr. Top. Dev. Biol*. 140, 391–427. doi: 10.1016/bs.ctdb.2020.05.001

Torres-Paz, J., Hyacinthe, C., Pierre, C., and Rétaux, S. (2018). Towards an integrated approach to understand Mexican cavefish evolution. *Biol. Lett*. 14:20180101. doi: 10.1098/rsbl.2018.0101

Torres-Paz, J., Leclercq, J., and Rétaux, S. (2019). Maternally-regulated gastrulation as a source of variation contributing to cavefish forebrain evolution. *ELife* 8:e50160. doi: 10.7554/eLife.50160

Varatharasan, N., Croll, R. P., and Franz-Odendaal, T. (2009). Taste bud development and patterning in sighted and blind morphs of *Astyanax mexicanus*. *Dev. Dyn*. 238, 3056–3064. doi: 10.1002/dvdy.22144

Vopalensky, P., Pralow, S., and Vastenhouw, N. L. (2018). Reduced expression of the nodal co-receptor oep causes loss of mesendodermal competence in zebrafish. *Development* 145:dev158832. doi: 10.1242/dev.158832

Williams, M. L. K., and Solnica-Krezel, L. (2020). Nodal and planar cell polarity signaling cooperate to regulate zebrafish convergence & extension gastrulation movements. *ELife* 9, 1–25. doi: 10.7554/eLife.54445

Woo, K., and Fraser, S. E. (1997). Specification of the zebrafish nervous system by nonaxial signals. *Science* 277, 254–257. doi: 10.1126/science.277.5323.254

Yamamoto, Y., and Jeffery, W. R. (2000). Central role for the lens in cave fish eye degeneration. *Science* 289, 631–633. doi: 10.1126/SCIENCE.289.5479.631

Yamamoto, Y., Stock, D. W., and Jeffery, W. R. (2004). Hedgehog signalling controls eye degeneration in blind cavefish. *Nature* 431, 844–847. doi: 10.1038/nature02906.1

Yoshizawa, M., Gorički, Š, Soares, D., and Jeffery, W. R. (2010). Evolution of a behavioral shift mediated by superficial neuromasts helps cavefish find food in darkness. *Curr. Biol*. 20, 1631–1636. doi: 10.1016/j.cub.2010.07.017

Yoshizawa, M., Hixon, E., and Jeffery, W. R. (2018). Neural crest transplantation reveals key roles in the evolution of cavefish development. *Integr. Comp. Biol*. 58, 411–420. doi: 10.1093/icb/icy006

Zuber, M. E., Gestri, G., Viczian, A. S., Barsacchi, G., and Harris, W. A. (2003). Specification of the vertebrate eye by a network of eye field transcription factors. *Development* 130, 5155–5167. doi: 10.1242/dev.00723

Conflict of Interest: 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 © 2021 Torres-Paz and Rétaux. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 April 2021
doi: 10.3389/fcell.2021.668025

The Ontology of the Amphioxus Anatomy and Life Cycle (AMPHX)

Stephanie Bertrand[1]*, João E. Carvalho[2†], Delphine Dauga[3], Nicolas Matentzoglu[4], Vladimir Daric[1], Jr-Kai Yu[5,6], Michael Schubert[2] and Hector Escrivá[1]*

[1] CNRS, Biologie Intégrative des Organismes Marins, Sorbonne Université, Paris, France, [2] CNRS, Laboratoire de Biologie du Développement de Villefranche-sur-Mer, Institut de la Mer de Villefranche, Sorbonne Université, Paris, France, [3] Bioself Communication, Marseille, France, [4] European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom, [5] Institute of Cellular and Organismic Biology, Academia Sinica, Taipei City, Taiwan, [6] Marine Research Station, Institute of Cellular and Organismic Biology, Academia Sinica, Yilan, Taiwan

OPEN ACCESS

Edited by:
*Maria Ina Arnone,
Department of Biology and Evolution of Marine Organisms, Zoological Station Anton Dohrn, Italy*

Reviewed by:
*Iryna Kozmikova,
Institute of Molecular Genetics (ASCR), Czechia
Filomena Ristoratore,
Department of Biology and Evolution of Marine Organisms, Zoological Station Anton Dohrn, Italy*

***Correspondence:**
*Stephanie Bertrand
stephanie.bertrand@obs-banyuls.fr
Hector Escrivá
hector.escriva@obs-banyuls.fr*

†Present address:
*João E. Carvalho,
CNRS, INSERM, Institute for Research on Cancer and Aging, Nice, Université Côte d'Azur, Nice, France*

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *15 February 2021*
Accepted: *31 March 2021*
Published: *26 April 2021*

Citation:
Bertrand S, Carvalho JE, Dauga D, Matentzoglu N, Daric V, Yu J-K, Schubert M and Escrivá H (2021) The Ontology of the Amphioxus Anatomy and Life Cycle (AMPHX). Front. Cell Dev. Biol. 9:668025. doi: 10.3389/fcell.2021.668025

An ontology is a computable representation of the different parts of an organism and its different developmental stages as well as the relationships between them. The ontology of model organisms is therefore a fundamental tool for a multitude of bioinformatics and comparative analyses. The cephalochordate amphioxus is a marine animal representing the earliest diverging evolutionary lineage of chordates. Furthermore, its morphology, its anatomy and its genome can be considered as prototypes of the chordate phylum. For these reasons, amphioxus is a very important animal model for evolutionary developmental biology studies aimed at understanding the origin and diversification of vertebrates. Here, we have constructed an amphioxus ontology (AMPHX) which combines anatomical and developmental terms and includes the relationships between these terms. AMPHX will be used to annotate amphioxus gene expression patterns as well as phenotypes. We encourage the scientific community to adopt this amphioxus ontology and send recommendations for future updates and improvements.

Keywords: amphioxus (lancelet), ontology (ontologie), cephalochordates, evodevo model organisms, life cycle, anatomy

INTRODUCTION

Cephalochordates (i.e., amphioxus) are a group of benthic marine filter feeding animals which live buried in the sand of shallow coastal environments in most temperate and tropical seas. Although there are only about thirty species and three genera (*Branchiostoma*, *Epigonichthys*, and *Asymmetron*), this chordate subphylum attracts the attention of numerous researchers because of its key phylogenetic position, representing the earliest diverging evolutionary lineage within chordates, and for its extraordinary morphological, anatomical, and genomic conservation with the last common ancestor of all chordates, including vertebrates (Bertrand and Escriva, 2011; Escriva, 2018). Some of these conserved morphological characters include a dorsal hollow neural tube and a dorsal notochord, pharyngeal slits, segmented muscles, and gonads as well as organs homologous to those of vertebrates, such as the pronephric kidney or an endostyle. However, some vertebrate-specific structures or organs are absent from amphioxus. These include paired sensory organs (eyes and ears), limbs or migrating neural crest cells. At the genomic level, amphioxus is also vertebrate-like but simpler, since their genomes did not experience the whole genome duplications that occurred in the vertebrate lineage (Dehal and Boore, 2005; Putnam et al., 2008). In addition, the amphioxus genome also shows simplified 3D structure and genetic regulation when compared to vertebrate genomes (Acemel et al., 2016; Marletaz et al., 2018).

All amphioxus species are gonochoric, and reproduce by external fertilization. Amphioxus males and females release their gametes in the water column during the spawning season, which, depending on the species, spans 3–6 months every year, usually during spring and summer time (Bertrand and Escriva, 2011). Amphioxus have a typical bentho-pelagic life cycle (**Figure 1**). An embryonic period is followed by a larval period, which ends when the larva undergoes metamorphosis, which, depending on the species, takes place from a few weeks to several months after fertilization. While amphioxus embryos and larva are planktonic, the juveniles emerging after metamorphosis are benthic and live buried into the substrate. The morphology of the post-metamorphic juvenile is identical to that of the adult. However, juveniles need to grow for a period of time, which can range from a few weeks in tropical species to several years in temperate species, before they start developing gonads and thus become sexually mature adults (Bertrand and Escriva, 2011).

Due to their phylogenetic position and extreme conservation, many laboratories around the world are interested in the study of the mechanisms controlling the embryonic development of amphioxus as well as in other aspects of its biology, such as regenerative capacities or its immune system (Somorjai et al., 2012; Yuan et al., 2015; Holland and Somorjai, 2020). These studies are mainly carried out on three or four species of the genus *Branchiostoma*, but recently also on a species of the genus *Asymmetron* (Holland and Holland, 2010). For this reason, and in order to standardize the results obtained in different species, we established a universal staging system (Carvalho et al., 2021) as well as an ontology of the development and anatomy of this chordate subphylum. This ontology, AMPHX, is the first one developed for a cephalochordate and follows the example of ontologies previously established for other chordate species (Segerdell et al., 2008; Manni et al., 2014; Van Slyke et al., 2014; Hotta et al., 2020). The AMPHX ontology has been conceived as an open and implementable automated retrieval system that can be integrated with biological information, such as gene expression or phenotypes. We hope that this ontology will be improved and updated whenever new studies on amphioxus appear in the literature.

MATERIALS AND METHODS

The different AMPHX terms, synonyms, definitions, and information on embryonic development and anatomical structures have been obtained from textbooks, journal articles, and scientific observations. This information has been compiled and formatted in two table files (**Supplementary Tables 1, 2**): one for life cycle and one for anatomy terms. The anatomy ontology was built following a bottom-up approach, starting from the cells and anatomical structures present in early stages of development, gradually introducing additional terms as new structures appear during embryogenesis, metamorphosis, and during juvenile and adult periods. We used the Webulous server[1] to create spreadsheets with an initial set of terms and relationships for the ontology. The resulting spreadsheet was integrated into an ontology development kit (ODK) workflow, using ROBOT (Jackson et al., 2019) templates, which we used for managing the transformation of our spreadsheets into the Web ontology language (OWL). We manage our ontology curation as well as collaborative and release workflows through GitHub[2].

RESULTS

Ontology Access

In the AMPHX ontology, we have described the anatomy and development, from the oocyte to the adult, of the cephalochordate amphioxus (i.e., the *Branchiostoma* genus). AMPHX can be downloaded from the AMPHX GitHub repository (see text footnote 2). The ontology was officially accepted into the OBO foundry and is listed in the OBO Foundry portal[3] (Smith et al., 2007) and the Ontobee database[4] (Ong et al., 2017). Users can also browse AMPHX at the ontology lookup service (OLS) at EMBL-EBI[5] (Côté et al., 2008; Jupp et al., 2015) and Bioportal[6] (Noy et al., 2009). In order for our data to be consistent with the FAIR principle[7], we submitted the AMPHX ontology to the FAIRsharing repository (Sansone et al., 2019), which approved our entry[8].

Current Content

The AMPHX ontology tree can be divided into two groups: a developmental entity, from AMPHX:0000001 to AMPHX:0000058 (**Supplementary Table 1**), and an anatomical entity, from AMPHX:1000001 to AMPHX:1000342 (**Supplementary Table 2**). In addition, there are three general terms from the cell ontology (CL[9]), corresponding to the following three major classes: cellular component (GO:0005575), cell (CL:0000000), and anatomical structure (UBERON:0000061), with a total number of 403 terms (**Figure 2**). The AMPHX ontology follows a structure-based subclass hierarchy including a partonomy relation between different terms ("part_of", BFO_0000050), a developmental hierarchy ("develops_from", RO_0002202), and a global developmental presence of a given term ("existence_starts_during" and "existence_ends_during", RO_0002488 and RO_0002492, respectively). The developmental hierarchy and presence terms are equivalent to the relation ontology[10] (RO). In the structure of AMPHX, each anatomical entity is thus defined by terms describing: being a type of, being a part of, developing from, and existing at (**Figure 3**).

---

[1] https://github.com/EBISPOT/webulous

[2] https://github.com/EBISPOT/amphx_ontology

[3] http://obofoundry.org/ontology/amphx.html

[4] http://www.ontobee.org/ontology/AMPHX

[5] https://www.ebi.ac.uk/ols/ontologies/amphx

[6] http://bioportal.bioontology.org/ontologies/AMPHX?p=summary

[7] https://en.wikipedia.org/wiki/FAIR_data

[8] https://fairsharing.org/bsg-s001582/

[9] http://purl.obolibrary.org/obo/cl.owl

[10] http://purl.obolibrary.org/obo/ro.owl

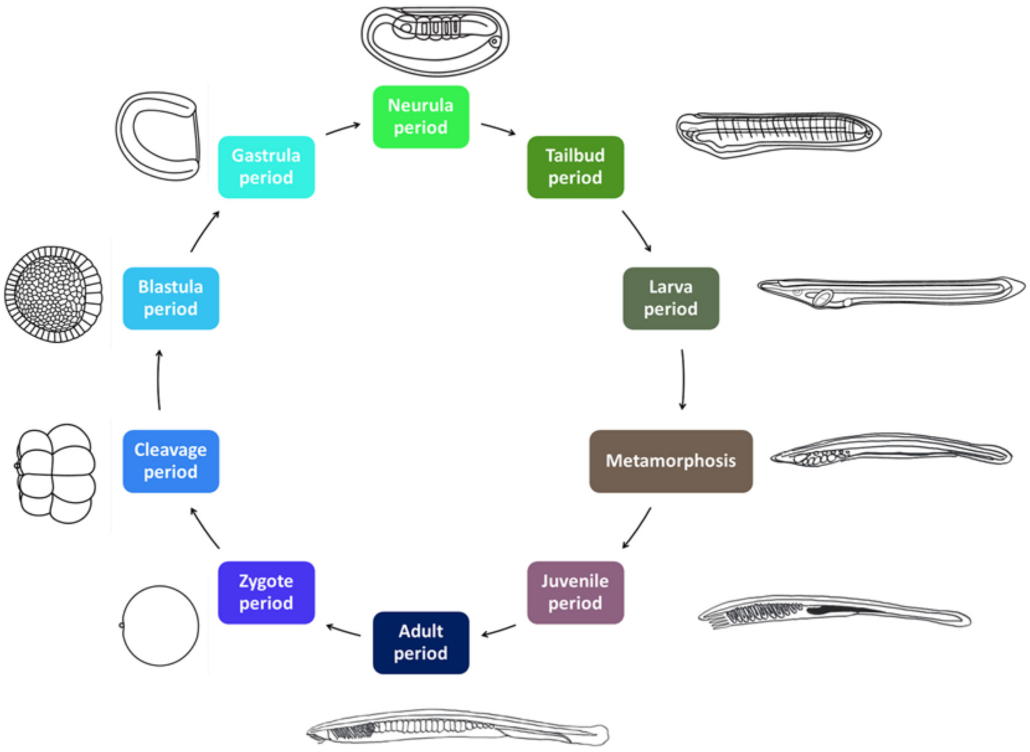


FIGURE 1 | The amphioxus life cycle. The ten developmental periods defined in AMPHX are indicated. Periods between zygote and metamorphosis are planktonic. The juvenile and adult periods are benthic. The size of the *Branchiostoma lanceolatum* embryos and larvae ranges from 80 to 100 μm in the zygote and early embryonic stages to several millimeters in the larval stages. Metamorphosis occurs in larvae with sizes of between 5 and 7 mm. Juveniles range in size from 5–7 mm to 3 cm and adults from 3 to 8 cm. Illustrations from zygote to larva periods have been adapted from Carvalho et al. (2021).

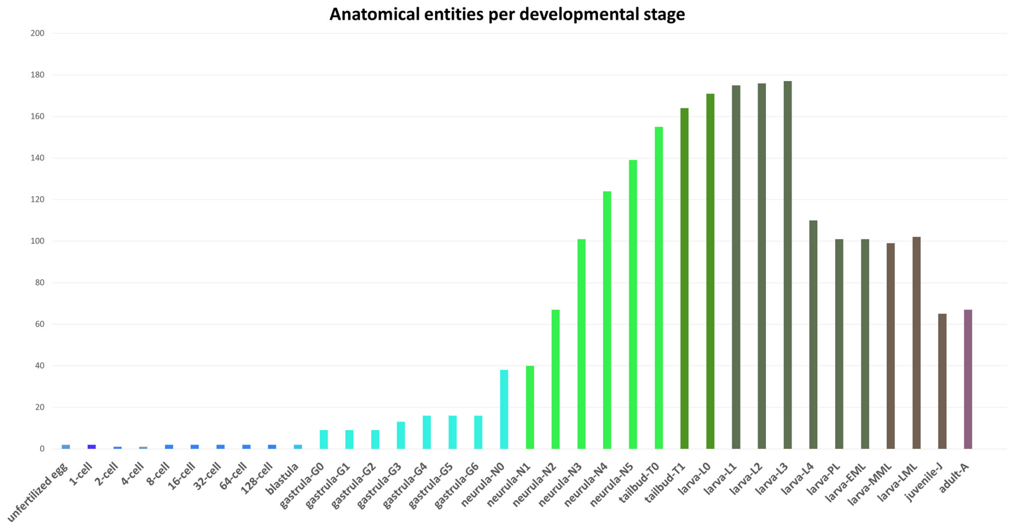


FIGURE 2 | Number of anatomical entities per developmental stage of *Branchiostoma lanceolatum*. The number of anatomical entities (Y-axis) in the unfertilized egg as well as at each developmental stage from the 1-cell stage to the adult stage (X-axis) is indicated.

One of the main aims of the AMPHX is to provide a queryable reference source for information about wild-type amphioxus development and anatomy. For this reason, we have added textual definitions to over 17% of the classes, based on scientific evidence derived from experimental observations. The overall number of these definitions will increase over time, with novel results and data that can be added to the AMPHX ontology.

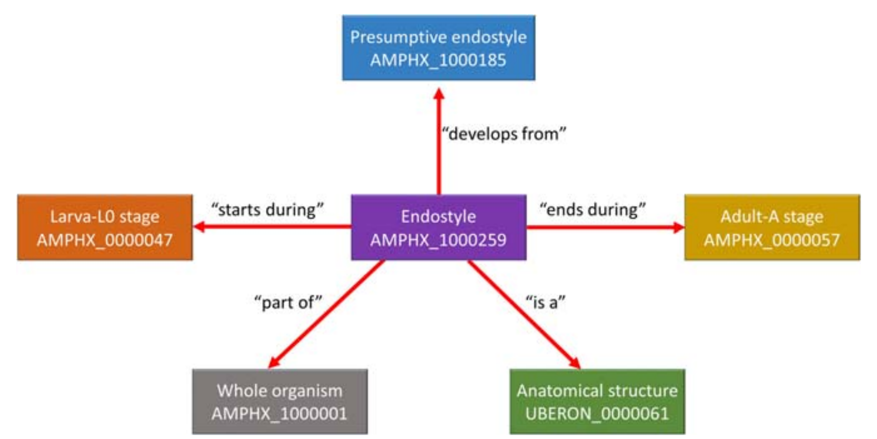


FIGURE 3 | Example for the representation of a structure-based subclass hierarchy in the amphioxus ontology (AMPHX). Using the amphioxus endostyle as an example, the partonomy relationships between different terms ("part_of", "develops_from", and "is a") as well as the developmental stages during which the structure is present ("starts during" and "ends during") are shown. The figure has been adapted from the Graph view utility proposed at OLS (Côté et al., 2008).

Integration of ontologies at successive stages was achieved *via* lineage links between mother and daughter territories. 323 such lineage links ("develops_from") were established, corresponding to 80% of the classes. This level of lineage information is vastly superior to the repertoires of ontologies of other model organisms, such as the *Caenorhabditis elegans* Gross Anatomy Ontology with 0,5% of classes with lineage information (Lee and Sternberg, 2003), the Zebrafish Anatomy and Development Ontology with 17% of classes with lineage information (Van Slyke et al., 2014) or the Xenopus Anatomy Ontology with 38% of classes with lineage information (Segerdell et al., 2008).

The AMPHX development ontology is based on *in vivo* observations [see Carvalho et al. (2021) for a morphological description of the different developmental stages] and is also inspired by the staging system proposed by Hirakow and Kajita (1990, 1991, 1994), which we have expanded by adding and redefining some developmental stages. AMPHX thus presents amphioxus development subdivided into ten different periods from zygote to adult (**Figure 1** and **Table 1**). We first defined the development ontology and subsequently added the anatomy ontology by carefully annotating anatomical structures for each developmental stage, based on information found in the literature. This allowed us not only to include all previously defined cell types, tissues, and organs but also to follow their differentiation through development. This led us to define 342 anatomical terms and assign a specific ID to each one of them.

The Development Ontology of *Branchiostoma* sp.

Embryonic Development From the Zygote to the Tailbud Stages

The early embryonic development is subdivided into four developmental periods: zygote period (AMPHX:0000008), cleavage period (AMPHX:0000009), blastula period (AMPHX:0000010), and gastrula period (AMPHX:0000011) (**Figure 1** and **Table 1**).

The zygote period comprises only one developmental stage, the 1-cell stage (AMPHX:0000022), which corresponds to the fertilized egg and ends with the first cell division. The cleavage period comprises the first cell divisions, which occur synchronously. This period includes seven developmental stages, from the 2-cell to the 128-cell stage (from AMPHX:0000023 to AMPHX:0000029). The first asynchronous cell divisions mark the start of the blastula period, which includes a single developmental stage, the blastula stage (AMPHX:0000031). The gastrula period includes seven developmental stages: the gastrula-G0 stage through the gastrula-G6 stage (AMPHX:0000032 to AMPHX:0000038). The gastrula-G0 stage corresponds to the initial flattening of the vegetal zone of the embryo. The gastrula-G1 stage corresponds to the stage at which the vegetal pole is flattened. During the gastrula-G2 stage, the vegetal pole is invaginating, but the mesendoderm does not touch the ectoderm. As soon as the mesendoderm touches the ectoderm, the embryo reaches the gastrula-G3 stage, also called cap-shaped gastrula. The gastrula-G4 stage corresponds to the cup-shaped gastrula. At the gastrula-G5 stage, the gastrula starts to elongate during blastopore closure, forming the vase-shaped gastrula. The final gastrula stage, the gastrula-G6 stage, corresponds to a bottle-shaped gastrula. At this stage, the neuroectoderm starts to flatten [for a morphological description of these stages see Carvalho et al. (2021)].

The gastrula period is followed by two additional periods of embryonic development, the neurula period (AMPHX:0000012), which immediately follows the gastrula period, and the tailbud period (AMPHX:0000013), which immediately follows the neurula period. Neurulation starts at the neurula-N0 stage (AMPHX:0000040) when the neural plate is visible, the blastopore is just about to close, and no somites have formed yet. The neurula stages following the neurula-N0 stage have been defined according to the number of somites that have developed in the embryo. Thus, the neurula-N1 stage embryo (AMPHX:0000041) has 1–3 somite pairs, and the neurula-N2 stage embryo (AMPHX:0000042) 4–5 somite pairs. The

TABLE 1 | Amphioxus developmental periods and stages from the zygote to the adult.

Period	Stage	ID	Description
Zygote period		AMPHX:0000008	
	1-cell stage	AMPHX:0000022	Fertilized egg before first cell division
Cleavage period		AMPHX:0000009	
	2-cell stage	AMPHX:0000023	
	4-cell stage	AMPHX:0000024	
	8-cell stage	AMPHX:0000025	
	16-cell stage	AMPHX:0000026	
	32-cell stage	AMPHX:0000027	
	64-cell stage	AMPHX:0000028	
	128-cell stage	AMPHX:0000029	
Blastula period		AMPHX:0000010	
	Blastula stage	AMPHX:0000031	Initiation of asynchronous cell division
Gastrula period		AMPHX:0000011	
	Gastrula-G0 stage	AMPHX:0000032	Initial flattening of the vegetal zone
	Gastrula-G1 stage	AMPHX:0000033	Flattened vegetal pole
	Gastrula-G2 stage	AMPHX:0000034	Invaginated vegetal pole, mesendoderm does not touch ectoderm
	Gastrula-G3 stage	AMPHX:0000035	Cap shaped, mesendoderm touches ectoderm
	Gastrula-G4 stage	AMPHX:0000036	Cup shaped
	Gastrula-G5 stage	AMPHX:0000037	Vase shaped, due to body elongation during blastopore closure
	Gastrula-G6 stage	AMPHX:0000038	Bottle shaped, flattening neuroectoderm
Neurula period		AMPHX:0000012	
	Neurula-N0 stage	AMPHX:0000040	Neural plate, just before blastopore closure
	Neurula-N1 stage	AMPHX:0000041	1–3 somite pairs
	Neurula-N2 stage	AMPHX:0000042	4–5 somite pairs, hatching
	Neurula-N3 stage	AMPHX:0000043	6–7 somite pairs
	Neurula-N4 stage	AMPHX:0000044	8–9 somite pairs, prior to schizocoelic somite formation
	Neurula-N5 stage	AMPHX:0000058	10–11 somite pairs
Tailbud period		AMPHX:0000013	
	Tailbud-T0 stage	AMPHX:0000045	12 somite pairs, tailbud shape, and enlarged pharyngeal region
	Tailbud-T1 stage	AMPHX:0000046	13 somite pairs, mouth and pre-oral pit anlagen, and first pigment spot
Larva period		AMPHX:0000015	
	Larva-L0 stage	AMPHX:0000047	Open mouth, no open gill slits
	Larva-L1 stage	AMPHX:0000048	Open mouth, 1 open gill slit
	Larva-L2 stage	AMPHX:0000049	Open mouth, 2 open gill slits
	Larva-L3 stage	AMPHX:0000050	Open mouth, 3 open gill slits
	Larva-L4 stage	AMPHX:0000051	Open mouth, 4–13 gill slits
	Larva-PL stage	AMPHX:0000052	Open mouth, 14–15 gill slits
Metamorphosis		AMPHX:0000016	
	Larva-EML stage	AMPHX:0000053	Metapleural folds start to grow
	Larva-MML stage	AMPHX:0000054	Gill slit row starts to duplicate, hepatic cecum starts to develop
	Larva-LML stage	AMPHX:0000055	Mouth migration to frontal position
Juvenile period		AMPHX:0000018	
	Juvenile-J stage	AMPHX:0000056	Adult-like morphology, length of less than 3 cm
Adult period		AMPHX:0000020	
	Adult-A stage	AMPHX:0000057	Adult-like morphology, length of more than 3 cm

A short description of the stages is included. The descriptions have been made using Branchiostoma lanceolatum. This table needs to be adapted slightly when applied to other species (especially in the late neurula and premetamorphic stages, as the number of somites and/or gill slits, the main bases for defining the stages, varies between species).

neurula-N2 stage is further the stage during which the embryo hatches. The last enterocoelic somites form during the neurula-N3 (AMPHX:0000043) and -N4 (AMPHX:0000044) stages characterized by, respectively, 6–7, and 8–9 somite pairs. The neurula-N5 stage (AMPHX:0000058) has 10–11 somite pairs and is characterized by the formation of the first schyzozoelic somites. The neurula-N5 stage marks the end of the neurula period, which is followed by the tailbud period. This period is composed of only two developmental stages, the tailbud-T0 stage (AMPHX:0000045) and the tailbud-T1 stage (AMPHX:0000046).

During the tailbud-T0 stage, the twelfth somite pair is formed, and the embryo adopts a tailbud shape with an enlarged pharyngeal region. During the tailbud-T1 stage, the thirteenth somite pair is formed and the mouth and pre-oral pit anlagen are formed as well as the first pigment spot.

Post-embryonic Development From the Larva to the Adult

Embryonic development in amphioxus is considered to be completed as soon as the mouth opens, which is when the larva period commences (AMPHX:0000015). We have defined most of the larval stages according to the number of formed pharyngeal slits. However, from the moment when the first three pharyngeal slits have formed and until the initiation of metamorphosis, there are only very few anatomical changes in the larva (except for the successive addition of new pharyngeal slits). We have thus defined only a single developmental stage for amphioxus larvae with four to thirteen gill slits. In total, the larva period comprises six developmental stages, starting with larva-L0 stage (AMPHX:0000047), which is characterized by the opening of the mouth. The larva-L0 stage is followed by the larva-L1 (AMPHX:0000048), -L2 (AMPHX:0000049), and -L3 (AMPHX:0000050) stages, during which, respectively, the first, second, and third pharyngeal slits appear. Subsequently, during the larva-L4 stage (AMPHX:0000051), new pharyngeal slits are added sequentially until the thirteenth slit. This stage is followed by the final larva stage, the larva-PL stage (AMPHX:0000052). This premetamorphic larva is characterized, in *Branchiostoma lanceolatum,* by 14–15 pharyngeal slits.

Metamorphosis in amphioxus is an extremely complex process where new structures appear and others are reorganized (Paris et al., 2008). We have divided the metamorphosis period (AMPHX:0000016) into three stages, based on the most drastic anatomical changes. The first anatomical change, which is initiated at the beginning of metamorphosis, is the appearance of the metapleural folds on both sides of the body of the early metamorphic larva. This appearance of the metapleural folds marks the larva-EML stage (AMPHX:0000053). The subsequent mid-metamorphic larva, referred to as the larva-MML stage (AMPHX:0000054), is characterized by the duplication of the row of pharyngeal slits and the formation of the hepatic cecum. The metamorphic process is completed when the frontal migration of the mouth has occurred in the late-metamorphic larva at the larva-LML stage (AMPHX:0000055).

Following metamorphosis, the juvenile period (AMPHX:0000018) commences. This period comprises only one developmental stage, the juvenile-J stage (AMPHX:0000056). Juveniles show an adult-like morphology and anatomy but lack gonads. In *B. lanceolatum*, the presence of mature gonads has never been observed in animals that have not reached a length of at least 3 cm. Therefore, we considered that the juvenile-J stage includes animals that have completed metamorphosis, but have not reached a length of 3 cm.

The last developmental period of the amphioxus life cycle is the adult period (AMPHX:0000020), which only contains one stage, the adult-A stage (AMPHX:0000057), and refers to post-metamorphic animals with a size of over 3 cm.

In addition to these different developmental periods and stages, we have included in AMPHX another non-developmental period, the gametogenesis period (AMPHX:0000001), which includes both oogenesis and spermatogenesis (AMPHX:0000002 and AMPHX:0000003, respectively) as well as a non-developmental stage, the gamete stage (AMPHX:0000004). This gamete stage is composed of the unfertilized egg stage (AMPHX:0000005) and the spermatozoid stage (AMPHX:0000006).

The Amphioxus Anatomy Ontology

The AMPHX anatomy ontology uses three high level terms from the cell ontology (CL, see text footnote 9): cell (CL:0000000), cellular component (GO:0005575), and anatomical structure (UBERON:0000061). Moreover, we have defined amphioxus-specific terms to define the type of a structure (i.e., "is a"). These terms include for example "whole organism" (AMPHX:1000001), "embryonic territory" (AMPHX:1000005), and "tissue" (AMPHX:1000004). When homologous terms exist in other ontologies, we have included these terms as a subclass ("similar to"). For example, the term "blastomere" (AMPHX:1000002) is similar to the NCI Thesaurus OBO Edition term NCIT:C12518, defined as "a cell formed by cleavage division during embryogenesis", the term "zygote" (AMPHX:1000013) is similar to the identical term in the Spider Ontology SPD:0000786 and is defined as "diploid cell produced by the fusion of sperm cell nucleus and egg", and the term "presumptive neural plate" (AMPHX:1000041) is similar to the identical term in the Uber-anatomy ontology (Mungall et al., 2012) UBERON:0007284 and is defined as "a presumptive structure that has the potential to develop into a neural plate".

All anatomical entities annotated in the AMPHX ontology were analyzed during ontogeny of *Branchiostoma* sp., stage by stage, from the unfertilized egg to the adult. The early developmental stages contain a lower number of anatomical entities, which increase as development proceeds (**Figure 2**). Thus, the number of entities associated with the developmental stages between the zygote period and the blastula period is 20, with the gastrula period 88, with the neurula period 509, with the tailbud period 319, and with the larva period 1,212. The number of associated entities decreases in the juvenile and adult periods. This effect can be explained by a relative paucity of information, in the scientific literature, on anatomical structures in amphioxus juveniles and adults.

Different relationships of a given AMPHX term can be analyzed using the graphics tool proposed in OLS (see text footnote 5) (Côté et al., 2008; Jupp et al., 2015). For example, an anatomical entity, such as the endostyle (AMPHX:1000259), which is similar to the term UBERON:0006870, defined as "a longitudinal ciliated groove on the ventral wall of the pharynx which produces mucus to gather food particles", can be represented as a function of its hierarchical relationships, including "is a", "develops from," and "part of", as well as of its development ("existence starts during", "existence ends during") (**Figure 3**). This representation can be expanded to show its complete development from the earliest cell types, tissues or organs (**Figure 4**). Explanations and a visual demonstration of

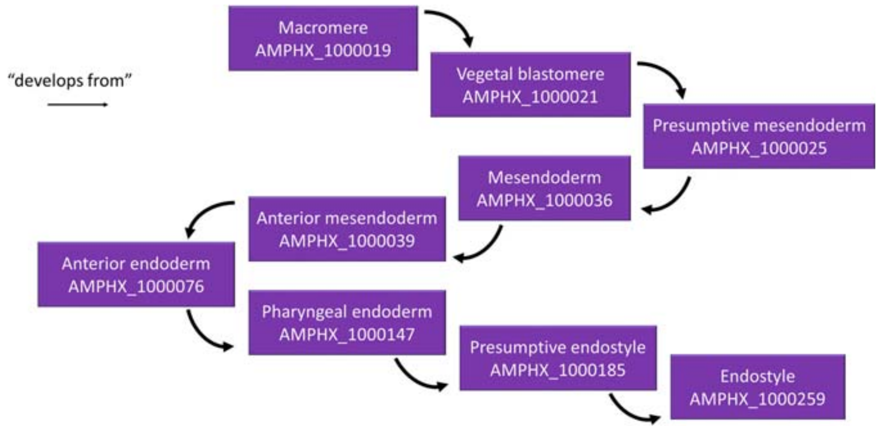

"develops from"
Macromere AMPHX_1000019
Vegetal blastomere AMPHX_1000021
Presumptive mesendoderm AMPHX_1000025
Mesendoderm AMPHX_1000036
Anterior mesendoderm AMPHX_1000039
Anterior endoderm AMPHX_1000076
Pharyngeal endoderm AMPHX_1000147
Presumptive endostyle AMPHX_1000185
Endostyle AMPHX_1000259

FIGURE 4 | Example for the hierarchical relationships in the AMPHX ontology. Using the amphioxus endostyle as an example, the anatomical entities giving rise to the structure ("develops from") are indicated, starting with the macromere in the 8-cell stage embryo. The figure has been adapted from the Graph view utility proposed at OLS (Côté et al., 2008).

the use of OLS, and in particular the use of the graphics tool, can be viewed on Youtube[11].

DISCUSSION

The AMPHX ontology has been designed to describe the anatomy and the development of amphioxus during its whole life cycle according to the principles of the open biomedical ontologies (OBO) Foundry (Smith et al., 2007). The AMPHX ontology will allow users to search and identify both developmental stages and anatomical structures and to follow the complete ontogeny of organs or structures of interest from the zygote to the adult. While the terms defining the development life cycle stages in AMPHX (from AMPHX:0000001 to AMPHX:00000058) can be directly applied to all amphioxus genera (i.e., *Branchiostoma*, *Epigonichthys*, and *Asymmetron*), the terms defining the anatomy (from AMPHX:1000001 to AMPHX:1000342) have been obtained from and adapted for the genus *Branchiostoma*. This implies that, although most of the terms can also be applied to the two other genera, it will be necessary to update AMPHX in the future to include terms specific for the *Epigonichthys* and *Asymmetron* genera.

The final objective of AMPHX is to allow the research community to use this partonomic definition of anatomical and developmental entities to assign gene expression patterns and phenotypic descriptions to each structure and developmental stage. A future objective of the AMPHX ontology should be its use as an interoperable ontology allowing comparisons between different amphioxus species and between amphioxus and other metazoan groups, particularly vertebrates. In order to perform such interspecies comparisons, homologous anatomical structures, and developmental stages have to be defined between the species being compared, and synonyms between the AMPHX ontology and other ontologies thus have to be established. However, these comparisons and the definition of homologous structures are not straightforward. While comparisons among different amphioxus species and genera are certainly feasible due to the high degree of anatomical and developmental conservation within the cephalochordate subphylum, such comparisons are significantly more complicated between amphioxus and members of other metazoan clades. The concept of biological homology, defined as sharing a common ancestor, is a binary relationship (a structure is or is not homologous to another structure). However, in an ontology it is necessary to consider relationships between cell types, tissues, and organs as a continuous rather than an all-or-nothing relationship (Wagner, 1989). As such, biological homology is unfortunately an impractical concept for cross-ontology comparisons. Ontology comparisons thus usually identify synonyms, rather than homologs. In the AMPHX ontology, we have been conservative with the definition of synonyms. We only used existing terms from other ontologies as a synonym of a term in AMPHX when the homology of the compared terms is generally accepted by the scientific community. For example, the amphioxus gastrula period (AMPHX:0000011) is homologous to the gastrula stage in UBERON (UBERON:0000109), which is defined as: "*A stage defined by complex and coordinated series of cellular movements that occurs at the end of cleavage during embryonic development of most animals. The details of gastrulation vary from species to species, but usually result in the formation of the three primary germ layers, ectoderm, mesoderm, and endoderm*". Given that this UBERON definition allows the inclusion of slight differences in the gastrulation process between different animals, the use of this term as a synonym to the AMPHX gastrulation period was validated. Another example for a synonymous term is the amphioxus neural plate (AMPHX:1000048), which corresponds to the identical term in UBERON (UBERON:0003075), which is defined as: "*A region of embryonic ectodermal cells that lies directly above the notochord. During neurulation, they change shape and produce an infolding of the neural plate (the neural fold) that then seals to form the neural tube.*

---

[11] https://www.youtube.com/watch?v=p8XF7yJvW4E

The earliest recognizable dorsal ectodermal primordium of the central nervous system present near the end of gastrulation before infolding to form the neural keel; consists of a thickened pseudostratified epithelium."

Having been conservative with the definition of synonyms, some structures that share a common ancestor and are thus homologous have not been assigned synonyms in the AMPHX ontology. This is the case, for example, for amphioxus and vertebrate somites. Given that the derivatives of amphioxus and vertebrate somites are not identical (Gilbert, 2003; Mansfield et al., 2015; Aldea et al., 2019), the term somite is only partially synonymous in the context of a cross-ontology comparison between amphioxus and vertebrates. We thus preferred not to define somites as synonyms in AMPHX. This highlights an important aspect of the AMPHX ontology, its potential for being amended and improved by the community. In particular, the definitions of cell types, tissues, and organs in juveniles and adults are lagging behind those of embryos and larvae. The number of structures in juveniles and adults is currently considerably lower than that for the embryonic and larval stages. This seems counterintuitive, as an adult should possess a significantly higher number of anatomical structures than an embryo. This bias is due to the fact that the list of terms in the AMPHX ontology is based mostly on the existing literature, which is heavily focused toward studies of embryo and larval stages. To expand further the AMPHX ontology, it would thus be interesting to focus more attention on the characterization of the ultrastructure of amphioxus juveniles and adults.

In conclusion, we encourage the amphioxus community and other scientific communities interested in animal evolution to use the AMPHX ontology. Input and support from users will be crucial to update and improve AMPHX by adding and redefining terms as new data become available. Of particular immediate importance will be the addition of information for the underrepresented juvenile and adult stages. Data from single-cell studies could be very useful for the refinement of AMPHX through the inclusion of new cell types, and a more detailed definition of the cell types and tissues that constitute amphioxus at any given stage.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HE and SB created the initial version of AMPHX. All the authors corrected and updated the initial version of AMPHX. DD, VD, and NM transformed the ontology into OWL and released workflows through GitHub. HE wrote the manuscript. All authors read and approved the manuscript.

FUNDING

The ontology creation and development has been directly funded by H2020-INFRADEV-4-2014-2015 RIA – Coordinated Research Infrastructures Building Enduring Life Science Services (CORBEL). HE and SB were supported by the "Agence Nationale de la Recheche" under the Grants ANR-19-CE13-0011-01 and ANR-16-CE12-0008-01. J-KY was supported by the Ministry of Science and Technology, Taiwan, under the Grant MOST-105-2628-B-001-003-MY3. JE was a FCT doctoral fellow (SFRH/BD/86878/2012) and is currently supported by a FRM fellowship (SPF20170938703). MS, HE, and SB were funded by the CNRS.

ACKNOWLEDGMENTS

This work benefited from access to the Observatoire Oceanologique de Banyuls sur Mer (OOB), an EMBRC-France and EMBRC-ERIC France/OOB member. Support was provided by the European project Assemble Plus (H2020-INFRAIA-1-2016–2017; Grant No. 730984). This manuscript is a back-to-back submission with Carvalho et al. (2021).

SUPPLEMENTARY MATERIALS

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.668025/full#supplementary-material

REFERENCES

Acemel, R. D., Tena, J. J., Irastorza-Azcarate, I., Marlétaz, F., Gómez-Marín, C., de la Calle-Mustienes, E., et al. (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat. Genet.* 48, 336–341. doi: 10.1038/ng.3497

Aldea, D., Subirana, L., Keime, C., Meister, L., Maeso, I., Marcellini, S., et al. (2019). Genetic regulation of amphioxus somitogenesis informs the evolution of the vertebrate head mesoderm. *Nat. Ecol. Evol.* 3, 1233–1240. doi: 10.1038/s41559-019-0933-z

Bertrand, S., and Escriva, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Carvalho, J. E., Lahaye, F., Yong, L. W., Croce, J. C., Escrivá, H., Yu, J.-K., et al. (2021). An updated staging system for cephalochordate development: one table suits them all. *Front. Cell Dev. Biol.* 9:668006. doi: 10.3389/fcell.2021.668006

Côté, R. G., Jones, P., Martens, L., Apweiler, R., and Hermjakob, H. (2008). The ontology lookup service: more data and better tools for controlled vocabulary queries. *Nucleic Acids Res.* 36, W372–W376.

Dehal, P., and Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:e314. doi: 10.1371/journal.pbio.0030314

Escriva, H. (2018). My favorite animal, amphioxus: unparalleled for studying early vertebrate evolution. *Bioessays* 40:e1800130.

Gilbert, S. F. (2003). *Developmental Biology*, 7th Edn. Sunderland, MA: Sinauer Associates, Inc.

Hirakow, R., and Kajita, N. (1990). An electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: cleavage. *J. Morphol.* 203, 331–344. doi: 10.1002/jmor.1052030308

Hirakow, R., and Kajita, N. (1991). Electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: the gastrula. *J. Morphol.* 207, 37–52. doi: 10.1002/jmor.1052070106

Hirakow, R., and Kajita, N. (1994). Electron microscopic study of the development of Amphioxus, *Branchiostoma belcheri tsingtauense*: the neurula and larva. *Kaibogaku Zasshi* 69, 1–13.

Holland, N. D., and Holland, L. Z. (2010). Laboratory spawning and development of the Bahama lancelet, *Asymmetron lucayanum* (*Cephalochordata*): fertilization through feeding larvae. *Biol. Bull.* 219, 132–141. doi: 10.1086/bblv219n2p132

Holland, N. D., and Somorjai, I. M. L. (2020). Tail regeneration in a cephalochordate, the Bahamas lancelet, *Asymmetron lucayanum*. *J. Morphol.* 282, 217–229. doi: 10.1002/jmor.21297

Hotta, K., Dauga, D., and Manni, L. (2020). The ontology of the anatomy and development of the solitary ascidian *Ciona*: the swimming larva and its metamorphosis. *Sci. Rep.* 10:17916.

Jackson, R. C., Balhoff, J. P., Douglass, E., Harris, N. L., Mungall, C. J., and Overton, J. A. (2019). ROBOT: a tool for automating ontology workflows. *BMC Bioinformatics* 20:407.

Jupp, S., Burdett, T., Leroy, C., and Parkinson, H. (2015). "A new ontology lookup service at EMBL-EBI," in *Proceedings of the SWAT4LS International Confernce* (Cambridge: SWAT4LS), 118–119.

Lee, R. Y., and Sternberg, P. W. (2003). Building a cell and anatomy ontology of *Caenorhabditis elegans*. *Comp. Funct. Genomics* 4, 121–126. doi: 10.1002/cfg.248

Manni, L., Gasparini, F., Hotta, K., Ishizuka, K. J., Ricci, L., Tiozzo, S., et al. (2014). Ontology for the asexual development and anatomy of the colonial chordate *Botryllus schlosseri*. *PLoS One* 9:e96434. doi: 10.1371/journal.pone.0096434

Mansfield, J. H., Haller, E., Holland, N. D., and Brent, A. E. (2015). Development of somites and their derivatives in amphioxus, and implications for the evolution of vertebrate somites. *Evodevo* 6:21.

Marletaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564, 64–70.

Mungall, C. J., Torniai, C., Gkoutos, G. V., Lewis, S. E., and Haendel, M. A. (2012). Uberon, an integrative multi-species anatomy ontology. *Genome Biol.* 13:R5. doi: 10.1515/jib-2007-65

Noy, N. F., Shah, N. H., Whetzel, P. L., Dai, B., Dorf, M., Griffith, N., et al. (2009). BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.* 37, W170–W173.

Ong, E., Xiang, Z., Zhao, B., Liu, Y., Lin, Y., Zheng, J., et al. (2017). Ontobee: A linked ontology data server to support ontology term dereferencing, linkage, query and integration. *Nucleic Acids Res.* 45, D347–D352.

Paris, M., Escriva, H., Schubert, M., Brunet, F., Brtko, J., Ciesielski, F., et al. (2008). Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr. Biol.* 18, 825–830. doi: 10.1016/j.cub.2008.04.078

Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071.

Sansone, S. A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A. L., et al. (2019). FAIRsharing as a community approach to standards, repositories and policies. *Nat. Biotechnol.* 37, 358–367. doi: 10.1038/s41587-019-0080-8

Segerdell, E., Bowes, J. B., Pollet, N., and Vize, P. D. (2008). An ontology for *Xenopus* anatomy and development. *BMC Devel. Biol.* 8:92. doi: 10.1186/1471-213X-8-92

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., et al. (2007). The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat. Biotechnol.* 25, 1251–1255. doi: 10.1038/nbt1346

Somorjai, I. M., Somorjai, R. L., Garcia-Fernandez, J., and Escriva, H. (2012). Vertebrate-like regeneration in the invertebrate chordate amphioxus. *Proc. Natl. Acad. Sci. U.S.A.* 109, 517–522. doi: 10.1073/pnas.1100045109

Van Slyke, C. E., Bradford, Y. M., Westerfield, M., and Haendel, M. A. (2014). The zebrafish anatomy and stage ontologies: representing the anatomy and development of *Danio rerio*. *J. Biomed. Semantics* 5:12. doi: 10.1186/2041-1480-5-12

Wagner, G. P. (1989). The biological homology concept. *Annu. Rev. Ecol. Syst.* 20, 51–69. doi: 10.1146/annurev.es.20.110189.000411

Yuan, S., Ruan, J., Huang, S., Chen, S., and Xu, A. (2015). Amphioxus as a model for investigating evolution of the vertebrate immune system. *Dev. Comp. Immunol.* 48, 297–305. doi: 10.1016/j.dci.2014.05.004

Conflict of Interest: 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 IK declared a past collaboration with one of the authors J-KY to the handling editor.

Copyright © 2021 Bertrand, Carvalho, Dauga, Matentzoglu, Daric, Yu, Schubert and Escrivá. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fcell.2021.668006

An Updated Staging System for Cephalochordate Development: One Table Suits Them All

João E. Carvalho[1†], François Lahaye[1], Luok Wen Yong[2], Jenifer C. Croce[1], Hector Escrivá[3], Jr-Kai Yu[2,4] and Michael Schubert[1]*

[1] Laboratoire de Biologie du Développement de Villefranche-sur-Mer, Institut de la Mer de Villefranche, Sorbonne Université, CNRS, Villefranche-sur-Mer, France, [2] Institute of Cellular and Organismic Biology, Academia Sinica, Taipei, Taiwan, [3] Biologie Intégrative des Organismes Marins, Observatoire Océanologique, Sorbonne Université, CNRS, Banyuls-sur-Mer, France, [4] Marine Research Station, Institute of Cellular and Organismic Biology, Academia Sinica, Yilan, Taiwan

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
RIKEN Cluster for Pioneering Research (CPR), Japan

Reviewed by:
David Ellard Keith Ferrier,
University of St Andrews, United Kingdom
Guang Li,
Xiamen University, China

***Correspondence:**
Michael Schubert
michael.schubert@obs-vlfr.fr

†Present address:
João E. Carvalho,
Institute for Research on Cancer and Aging, Nice (IRCAN), CNRS, INSERM, Université Côte d'Azur, Nice, France

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *15 February 2021*
Accepted: *31 March 2021*
Published: *20 May 2021*

Citation:
Carvalho JE, Lahaye F, Yong LW, Croce JC, Escrivá H, Yu J-K and Schubert M (2021) An Updated Staging System for Cephalochordate Development: One Table Suits Them All. Front. Cell Dev. Biol. 9:668006. doi: 10.3389/fcell.2021.668006

Chordates are divided into three subphyla: Vertebrata, Tunicata, and Cephalochordata. Phylogenetically, the Cephalochordata, more commonly known as lancelets or amphioxus, constitute the sister group of Vertebrata and Tunicata. Lancelets are small, benthic, marine filter feeders, and their roughly three dozen described species are divided into three genera: *Branchiostoma*, *Epigonichthys*, and *Asymmetron*. Due to their phylogenetic position and their stereotypical chordate morphology and genome architecture, lancelets are key models for understanding the evolutionary history of chordates. Lancelets have thus been studied by generations of scientists, with the first descriptions of adult anatomy and developmental morphology dating back to the 19th century. Today, several different lancelet species are used as laboratory models, predominantly for developmental, molecular and genomic studies. Surprisingly, however, a universal staging system and an unambiguous nomenclature for developing lancelets have not yet been adopted by the scientific community. In this work, we characterized the development of the European lancelet (*Branchiostoma lanceolatum*) using confocal microscopy and compiled a streamlined developmental staging system, from fertilization through larval life, including an unambiguous stage nomenclature. By tracing growth curves of the European lancelet reared at different temperatures, we were able to show that our staging system permitted an easy conversion of any developmental time into a specific stage name. Furthermore, comparisons of embryos and larvae from the European lancelet (*B. lanceolatum*), the Florida lancelet (*Branchiostoma floridae*), two Asian lancelets (*Branchiostoma belcheri* and *Branchiostoma japonicum*), and the Bahamas lancelet (*Asymmetron lucayanum*) demonstrated that our staging system could readily be applied to other lancelet species. Although the detailed staging description was carried out on developing *B. lanceolatum*, the comparisons with other lancelet species thus strongly suggested that both staging and nomenclature are applicable to all extant lancelets. We conclude that this description of embryonic and larval development will be of great use for the scientific community and that it should be adopted as the new standard for defining and naming developing lancelets. More generally, we anticipate that this work will facilitate future studies comparing representatives from different chordate lineages.

Keywords: amphioxus, lancelet, *Branchiostoma lanceolatum*, *Branchiostoma floridae*, *Branchiostoma belcheri*, *Branchiostoma japonicum*, *Asymmetron lucayanum*, evolution and development

INTRODUCTION

The subphylum Cephalochordata comprises only a few dozen species of small, lancet-shaped filter-feeders (Bertrand and Escrivá, 2011; Holland, 2015). The Cephalochordata (commonly referred to as lancelets or amphioxus) belong to the chordate phylum and are the sister group to all other chordates (Tunicata and Vertebrata) (Bertrand and Escrivá, 2011; Holland, 2015). Due to this phylogenetic position and their slow evolutionary rate (Louis et al., 2012), lancelets are considered valuable proxies for the chordate ancestor, both at the anatomic and genomic levels (Bertrand and Escrivá, 2011; Holland, 2015). The subphylum Cephalochordata is subdivided into three genera: *Branchiostoma*, *Epigonichthys*, and *Asymmetron* (Poss and Boschung, 1996; Nishikawa, 2004; Zhang et al., 2006; Kon et al., 2007; Holland and Holland, 2010; Yue et al., 2014; Carvalho et al., 2017b; Subirana et al., 2020). Recent analyses of mitochondrial genomes suggested that the genus *Asymmetron* occupies the basal position and diverged from the *Epigonichthys/Branchiostoma* clade about 258-171 mya (million years ago) (Subirana et al., 2020). It was further proposed that the split of the *Epigonichthys* and *Branchiostoma* lineages occurred about 182-120 mya and that speciation within the genus *Branchiostoma*, between *Branchiostoma belcheri* and *Branchiostoma japonicum* versus *Branchiostoma floridae* and *Branchiostoma lanceolatum*, took place about 130-85 mya (Subirana et al., 2020).

The importance of lancelets for understanding chordate evolution has driven generations of scientists to study their embryos and larvae (Holland and Holland, 2017). An initial description of lancelet development was already performed in the 19th century, on *B. lanceolatum* material obtained in Naples, Italy (Kovalevsky, 1867). This work was subsequently completed, at the end of the 19th and the beginning of the 20th century, by a series of additional surveys on the same species (Hatschek, 1893; Cerfontaine, 1906; Conklin, 1932). More recently, in the early 1990s, the early development of *B. japonicum* (previously known as *Branchiostoma belcheri tsingtauense*) was the subject of a detailed characterization by electron microscopy (Hirakow and Kajita, 1990, 1991, 1994). A similar approach was used to characterize neurulae, larvae, and post-metamorphic specimens of *B. floridae* (Holland and Holland, 1992; Stokes and Holland, 1995). The most recent description of lancelet development was that of *Asymmetron lucayanum* embryos and larvae using differential interference contrast (DIC) microscopy (Holland and Holland, 2010; Holland et al., 2015). Taken together, these studies on species of the two most distantly related lancelet genera have revealed that the ontogeny of lancelets is a highly coordinated and conserved process. It is thus all the more surprising that there is currently no universal developmental staging system available for the members of this subphylum.

In the course of the last three decades, lancelets have become important models for addressing developmental processes from a molecular and genomic perspective (Bertrand and Escrivá, 2011; Acemel et al., 2016; Carvalho et al., 2017b; Marlétaz et al., 2018; Simakov et al., 2020). However, unlike for other developmental model organisms, such as zebrafish, the scientific community is using different lancelet species for their studies, with the choice being mainly dependent on the availability of animal resources (Carvalho et al., 2017b). Husbandry protocols have been established for at least five lancelet species (Carvalho et al., 2017b), but, due to the absence of a universal staging system, the nomenclature of embryos and larvae obtained with these protocols has become extremely confusing. While developing lancelets are often named in accordance with previous reports on the same species (Bertrand et al., 2011; Lu et al., 2012; Holland, 2015; Annona et al., 2017), it is also not uncommon to indicate the time after fertilization, usually measured in hours after fertilization (Fuentes et al., 2007; Bertrand and Escrivá, 2011). However, developmental speed is known to vary between lancelet species and to depend on the rearing temperature, which is not the same in each study (Fuentes et al., 2007; Bertrand and Escrivá, 2011). The absence of an unambiguous nomenclature for developing lancelets artificially complicates comparisons of results obtained in different species and sometimes even within the same species, for example, when two laboratories use incompatible staging styles (Bertrand et al., 2011; Pantzartzi et al., 2017). There is, therefore, an urgent need to establish an easy and systematic classification for embryonic and larval development that applies to different lancelet species.

To achieve this objective, we illustrated the development of *B. lanceolatum* using confocal microscopy and established growth curves at different temperatures based on the number of somites. We further compared embryos and larvae of *B. lanceolatum* with those of other lancelets. By applying and expanding the stage definitions of Hirakow and Kajita (1990, 1991, 1994) and Lu et al. (2012), we compiled a streamlined staging system of *B. lanceolatum* development, from fertilization through larval life, with an unambiguous stage nomenclature. Analyses of growth curves revealed that our staging system could be used to easily convert developmental times into unambiguous stage names, at any given rearing temperature. Furthermore, comparisons between *B. lanceolatum*, *B. floridae*, *B. belcheri*, *B. japonicum*, and *A. lucayanum* embryos and larvae demonstrated that the updated staging system could readily be applied to other lancelet species. We hope that the scientific community will adopt this universal developmental staging system for lancelets to facilitate the use of these fascinating animals as laboratory models.

MATERIALS AND METHODS

Animal Husbandry and *in vitro* Cultures

Ripe *B. lanceolatum* adults were collected by dredging in Argelès-sur-Mer, France, and retrieved from the sand by sieving. Animals were transported, quarantined and maintained in Villefranche-sur-Mer as previously described (Carvalho et al., 2017b). Spawning was induced by a 36-h thermal shock at 23°C (Fuentes et al., 2004). Sperm and oocytes were collected separately and fertilization was performed *in vitro*. *B. lanceolatum* embryos and larvae were raised in the dark at constant temperatures (16°C, 19°C, or 22°C) until the desired developmental stages, and larvae were fed daily with *Tisochrysis lutea* algae (Carvalho et al., 2017b).

Adult *B. floridae* were collected in Tampa Bay, FL, United States. Animals were maintained in the laboratory as previously described (Zhang et al., 2007; Yong et al., 2019). Gametes were obtained either by electric stimulation, heat shock, or spontaneous spawning (Holland and Yu, 2004; Ono et al., 2018). Embryos and larvae were cultured at constant temperatures (25°C or 30°C) until the desired stages, and larvae were fed daily with *Isochrysis sp.* algae.

Adult *B. belcheri* and *B. japonicum* were collected in Kinmen Island near Xiamen in southeastern China (Zhang et al., 2013). Animals were maintained as previously described (Zhang et al., 2007; Yong et al., 2019). Embryos were obtained through spontaneous spawning in the facility (Zhang et al., 2007). Embryos and larvae were cultured at a constant temperature (24°C for *B. belcheri* and 25°C for *B. japonicum*) until the desired stages, and larvae were fed daily with *Isochrysis sp.* algae.

Asymmetron lucayanum adults were collected in the lagoon between North and South Bimini, Bahamas. Embryos and larvae were obtained and subsequently cultured at a constant temperature (27°C) as previously described (Holland and Holland, 2010).

Differential Interference Contrast (DIC) Microscopy

Embryos and larvae used for observation and imaging by DIC microscopy were fixed in 4% PFA (paraformaldehyde) in MOPS buffer for 1 h at room temperature or overnight at 4°C. Embryos and larvae were subsequently washed twice in ice-cold 70% ethanol in DEPC-treated water and stored at −20°C until further use. Embryos and larvae were rehydrated in PBS buffer and mounted in PBS buffer or 80% glycerol for imaging.

DIC microscopy of *B. lanceolatum* embryos and larvae was performed using a Zeiss Axiophot microscope, equipped with an AxioCam ERc 5s camera (Carl Zeiss SAS, Marly-le-Roi, France). Images of *B. floridae*, *B. belcheri*, *B. japonicum*, and *A. lucayanum* embryos and larvae were acquired with a Zeiss Axio Imager A1 microscope, equipped with a AxioCam HRc CCD camera (Carl Zeiss SAS, Marly-le-Roi, France). For 64-cell, 128-cell, and blastula stages, multiple z-levels were taken manually. The z-stack images were processed with the Extended-Depth-of-Field plugin of the ImageJ software using default settings (Schneider et al., 2012), and panels were subsequently formatted with Adobe Photoshop CS6 (Adobe Inc., San Jose, CA, United States).

Fluorescent Staining and Immunohistochemistry

B. lanceolatum fertilized eggs as well as cleavage- and gastrula-stage embryos were stained using FM 4-64 lipophilic dye (Invitrogen, Cergy-Pontoise, France) at a final concentration of 10 μg/ml (stock solution of 20 μg/ml in artificial sea water mixed 1:1 with live embryos). The FM 4-64 lipophilic dye is a non-toxic vital dye commonly used to label plasma membranes and endocytic pathways (Sardet et al., 2011). Following the 10 min dye incubation, the embryos were fixed for 1 h at room temperature with freshly prepared 4% PFA in MOPS buffer (Yu and Holland, 2009). Embryos were washed twice in 70% ethanol and subsequently rehydrated in PBS buffer (Yu and Holland, 2009). Nuclear DNA staining was performed for 10 min at room temperature using Hoechst dye (Invitrogen, Cergy-Pontoise, France) at a final dilution of 1:5000. Embryos were mounted in PBS buffer and imaged within 3 h after staining with the FM 4-64 and Hoechst dyes.

For neurula, tailbud, and larva stages, the FM 4-64 lipophilic dye yielded unsatisfactory results. These stages were thus stained by immunohistochemistry using a primary antibody against aPKC (polarity protein atypical protein kinase C), which labels structures associated with cell membranes (Patalano et al., 2006; Prulière et al., 2011). For whole-mount immunohistochemistry *B. lanceolatum* embryos and larvae were fixed overnight at 4°C in freshly prepared ice-cold 4% PFA in MOPS buffer (Yu and Holland, 2009). Immunohistochemistry was performed as previously described (Zieger et al., 2018), using the primary antibody against aPKC (SC216, Santa Cruz Biotechnology, Dallas, TX, United States) at a final dilution of 1:100 and a secondary anti-mouse IgG-heavy and light chain antibody conjugated with Cy3^{TM} (A90-516C3, Bethyl Laboratories, Inc., Montgomery, TX, United States) at a final dilution of 1:200. Hoechst dye (Invitrogen, Cergy-Pontoise, France) at a final dilution of 1:5000 was used for nuclear DNA staining. Embryos were mounted in PBS buffer and subsequently imaged.

Imaging was systematically carried out on a Leica TCS SP8 confocal microscope, using a 20x objective (0.75 IMM HC PL APO CORR CS WD = 0.68 mm) (Leica Microsystems SAS, Nanterre, France). FM 4-64/DNA staining and aPKC/DNA staining scans were obtained sequentially. DNA, FM 4-64, and aPKC staining were excited using, respectively, 405 nm, 514 nm, and 552 nm lasers. Series of optical sections were taken at a z-step interval of 2 μm. The ImageJ software (Schneider et al., 2012) was subsequently used for image processing and to generate maximum as well as average projections. Adobe Photoshop CS6 (Adobe Inc., San Jose, CA, United States) was used to format larger panels requiring the reconstitution of partial images.

Growth Curves and *in situ* Hybridization

Developing *B. lanceolatum* embryos were reared at three different temperatures: 16°C, 19°C, and 22°C. At regular intervals, animals were collected and fixed for subsequent *in situ* hybridization analyses. A 874-bp fragment containing the complete coding sequence of the *B. lanceolatum mrf1* (*myogenic regulatory factor 1*) gene, a member of the *myoD* gene family (Schubert et al., 2003; Aase-Remedios et al., 2020), was amplified by PCR from cDNA and cloned into the pGEM-T Easy Vector (GenBank accession number of *B. lanceolatum mrf1*: MT452570). *In situ* hybridization experiments were carried out with a *mrf1*-specific antisense riboprobe as previously described (Yu and Holland, 2009; Carvalho et al., 2017c). Following *in situ* hybridization, embryos and larvae were mounted for DIC microscopy and imaged as described above.

Expression of the *mrf1* gene was used to visualize the somites and thus to obtain somite pair counts of embryos and larvae reared at different temperatures. The somite pair counts were used to define a training set of data points for each rearing temperature (16°C, 19°C, and 22°C), hence allowing the calculation of best natural logarithmic tendency curves using Microsoft Excel (Microsoft Corporation, Redmond, WA, United States). The curves were subsequently curated and used to define time intervals for each developmental stage (i.e., cleavage, blastula, gastrula, neurula, tailbud, and larva stages).

Data Availability

The complete imaging dataset used to generate the main and **Supplementary Figures** (including confocal images and stacks as well as DIC images) is accessible at: http://movincell.com/projects/14.

RESULTS

Branchiostoma lanceolatum Staging Series

Making use of the available *in vitro* culture protocols for developing lancelets (Carvalho et al., 2017b), the updated staging system was established using *B. lanceolatum* embryos and larvae. Prior to confocal imaging, embryos, and larvae were fixed at the desired stages and labeled with fluorescent probes marking cell membranes and nuclei, hence allowing detailed morphological analyses of individual developmental stages. In the following, each stage of the updated staging system will be presented and defined. The stage names are indicative of the developmental period and are in accordance with previous descriptions of lancelet development (Hirakow and Kajita, 1990, 1991, 1994; Lu et al., 2012) as well as with the recently developed ontology for the *Branchiostoma* genus, AMPHX (Bertrand et al., 2021).

Fertilization and Cleavage

Lancelets are gonochoric and reproduce by external fertilization. Under appropriate environmental conditions, gravid males and females, respectively, release mature spermatozoa and oocytes into the water column. Prior to spawning, the mature lancelet oocyte undergoes the first meiotic division with formation of the first polar body and it is subsequently arrested in the second meiotic metaphase (Holland and Onai, 2012). Following spawning, the second meiotic division of the oocyte is triggered by fertilization and is completed within 10 min. The second meiotic division leads to the formation of the second polar body and the migration of the maternal chromosomes to the animal pole, which is defined by the position of the polar body (**Figure 1A** and **Supplementary Figure 1A**; Holland and Holland, 1992). Independent of the entry point, the nucleus of the sperm first migrates to the vegetal half and only then joins the maternal chromosomes at the animal pole (Holland, 2015). Very soon after fertilization, a whorl composed of sheets of endoplasmic reticulum is further formed within the 1-cell stage. This whorl likely constitutes the germ plasm, since expression of germ cell markers, such as *nanos* and *vasa*, is associated with this structure (Wu et al., 2011). The 1-cell stage embryo is semi-opaque, due to the high quantity of granules uniformly distributed throughout the cell, and is surrounded by a membrane called the vitelline layer (Willey, 1894). As soon as fertilization occurs, the vitelline layer detaches from the 1-cell stage and expands, giving rise to the fertilization envelope (Holland and Holland, 1989). Cleavage, gastrulation, and the first stages of neurulation will occur within the fertilization envelope (Holland, 2015).

Lancelet cleavage is radial holoblastic, meaning that cleavage completely separates blastomeres and results in early stage embryos with radial symmetry along the animal-vegetal axis (Barresi and Gilbert, 2019). The first cleavage starts from the animal pole and gives rise to the 2-cell stage, which is composed of two identically shaped blastomeres (**Figure 1B**). When dissociated, each one of the first two blastomeres can give rise to a complete animal, but only one of the two blastomeres inherits the germ plasm (Holland and Onai, 2012). The second division is meridional and at a right angle to the first one, creating four blastomeres with approximately equal size, the 4-cell stage (**Figure 1C**). Individual blastomeres are not adhering very strongly at this stage, and their dissociation can lead to the formation of twins or even quadruplets (Holland and Onai, 2012). Cleavage continues by an equatorial division, creating four animal and four vegetal blastomeres at the 8-cell stage, with the former being smaller than the latter (**Figure 1D**). The blastomeres are held together by short microvilli and slender filopodial processes that bridge the space between adjacent blastomeres (insets in **Figures 1D,E**; Hirakow and Kajita, 1990). The 16-cell stage is the result of a meridional cleavage (**Figure 1E**), and the 32-cell stage of a subsequent equatorial cleavage of each blastomere (**Figure 1F**). At the 32-cell stage, the embryo is composed of a single layer of cells forming a central cavity called the blastocoel (**Supplementary Figure 1B**; Grassé, 1948; Hirakow and Kajita, 1990). The blastomeres will keep dividing regularly, giving rise to the 64-cell stage (**Figure 1G**) and then to the 128-cell stage (**Figure 1H**). The 8th cell division cycle, i.e., the transition from 128 cells to 256 cells, which we define as the B stage, is characterized by the initiation of asynchronous cell division within the embryo (Grassé, 1948; Hirakow and Kajita, 1990) and further marks the formation of the blastula (**Figure 1I**). The cells constituting

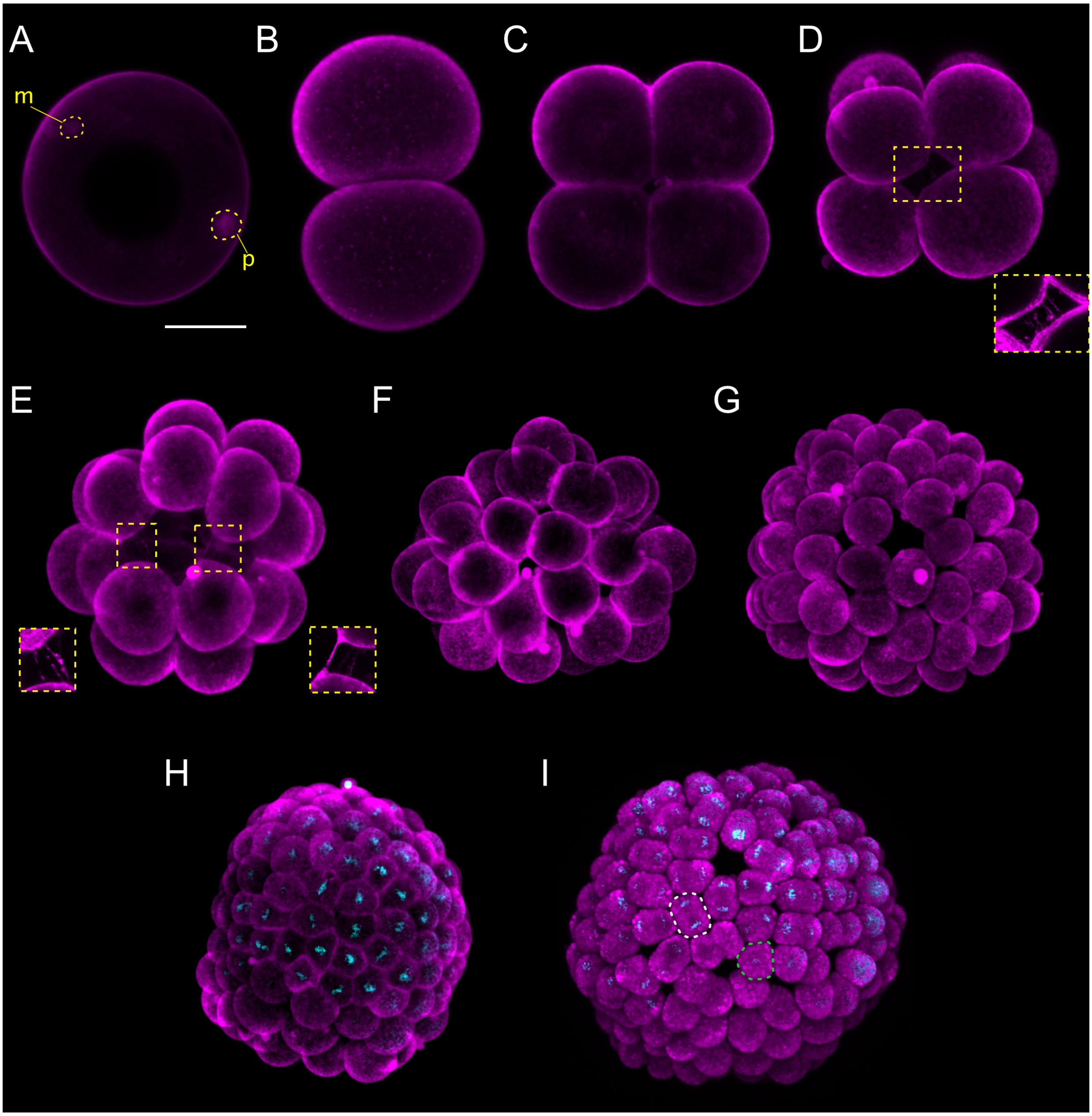


FIGURE 1 | *Branchiostoma lanceolatum* fertilization, cleavage, and blastula stages. Embryos are stained with the lipophilic dye FM 4-64 (magenta). **(B,C)** Animal pole views. **(D–I)** Animal pole is up. Maximum projections of confocal z-stacks of *B. lanceolatum* embryos at the **(A)** 1 cell-stage, **(B)** 2-cell stage, **(C)** 4-cell stage, **(D)** 8-cell stage, **(E)** 16-cell stage, **(F)** 32-cell stage, **(G)** 64-cell stage, **(H)** 128-cell stage, and **(I)** blastula B stage. Insets in **(D)** and **(E)** show slender filopodia between blastomeres. In **(H,I)**, Hoechst DNA staining (cyan) shows synchronous cell divisions at the 128-cell stage **(H)** and asynchronous cell divisions at the forming blastula B stage **(I)**, with a white dashed line highlighting a cell in telophase and a green dashed line highlighting a cell following cytokinesis. m, maternal DNA; p, paternal DNA. Scale bar: 50 μm.

the blastula will undergo two rounds of cell division before gastrulation is initiated.

Gastrulation

The cells forming the hollow blastula are not identical in shape and size. The vegetal blastula cells are larger and hence indicate where the initial flattening of the gastrula takes place at the G0 stage (**Figure 2A**; Willey, 1894; Holland, 2015). The vegetal side of the embryo will continue to flatten and bend inward at the G1 stage (**Figures 2B,B′**), hence forming a depression that marks the position of the blastopore. Thereafter, the vegetal tissue starts to invaginate into the blastocoel at the G2 stage (**Figures 2C,C′**; Hirakow and Kajita, 1991). The invaginating cells correspond to the presumptive endomesoderm, while the non-invaginating cells of the outer layer constitute the future general and neural ectoderm (Holland and Onai, 2012). As gastrulation proceeds with further cell divisions, the invaginating cells reduce the size of the blastocoelic cavity, ultimately leading, at the G3 stage, to a two-layered gastrula with an archenteron and a blastoporal lip. In this cap-shaped gastrula, the diameter of the blastopore is about half the size of the entire embryo (**Figures 2D,D′**; Hirakow and Kajita, 1991). Subsequent gastrulation movements result in an expansion of the cavity of the archenteron and in an almost complete loss of the blastocoelic cavity. This process leads to a narrowing of the blastoporal opening, which inflects the blastoporal lip, forming a cup-shaped gastrula at the G4 stage (**Figures 2E,E′**) and a vase-shaped gastrula at the G5 stage (**Figures 2F,F′**; Hirakow and Kajita, 1991). Starting at the G5 stage, differences between the dorsal and ventral sides of the embryo become discernable, with the dorsal side beginning to flatten (**Figures 2F,F′**; Willey, 1894). These differences become more pronounced at the G6 stage, as the size of the blastopore continues to decrease and the embryo continues to elongate (**Figures 2G,G′**). At this late gastrula stage, the embryo is bottle-shaped, and the blastopore starts to incline toward the dorsal side of the embryo, which is likely a synapomorphic trait of chordates, already present in their last common ancestor (Willey, 1894).

Expression patterns of marker genes have determined that, with the exception of the tissues located in the immediate vicinity of the blastopore, most of the gastrula is destined to become the anteriormost region of the lancelet larva. This includes the cerebral vesicle, the anteriormost somites, the pharynx with mouth and gill slits as well as the anterior section of the notochord (Holland and Onai, 2012). Transplantation experiments further demonstrated that the dorsal lip of the blastopore corresponds to a gastrulation organizer, similar or equivalent to the Spemann-Mangold organizer of vertebrates (Tung et al., 1961, 1962; Le Petillon et al., 2017).

Neurulation

Following gastrulation, ectodermal cells develop cilia (**Supplementary Figures 1C,C′**), and the embryo therefore starts to rotate within the fertilization envelope by ciliary movement (Hirakow and Kajita, 1991). Cilia are also present on the endomesodermal cells of the archenteron (Hirakow and Kajita, 1991), and these cilia have been shown to play a role in establishing left-right asymmetry (Zhu et al., 2020). At this point in development, the N0 stage, neurulation starts. The N0 stage embryo is unsegmented and shows a typical dipoblastic organization, with the ectoderm externally and the endomesoderm internally (**Figure 3A**). A small blastopore is still visible, and the dorsal ectoderm, destined to become the neuroectoderm, is flat with a shallow longitudinal groove (**Figure 3A**). The subsequent N1 stage is characterized by the establishment of the first somites (somite pairs 1 through 3) (**Figures 3B,B′**). The mesoderm, located dorsally within the endomesoderm, forms three folds: one medially that will develop into the notochord and two laterally that will give rise to the anterior somite pairs (**Supplementary Figure 1C′**). At the N1 stage, the somites start pinching off in an anterior to posterior sequence. At the same stage, the dorsal non-neural ectoderm starts to detach from the edges of the neural plate (Hirakow and Kajita, 1994). Following their detachment, the ectodermal cells will migrate over the neural plate using lamellipodia and fuse at the dorsal midline (Holland et al., 1996). At the end of this process, the neural plate will be completely covered by non-neural ectoderm, and the neuropore will have been formed anteriorly (**Supplementary Figure 1C**; Hatschek, 1881, 1893; Holland and Onai, 2012).

As neurulation proceeds, the archenteron is no longer in contact with the exterior, but still communicates with the forming neural tube: the blastopore is incorporated into the neurenteric canal, which connects the neural tube with the archenteron (**Supplementary Figures 1D,E,E′,E′′**), which becomes the presumptive gastric cavity (Willey, 1894). The embryo keeps elongating by the addition of new somites, reaching 4–5 somite pairs at the N2 stage (**Figures 3C,C′** and **Supplementary Figure 1E′**). At this stage, the embryo hatches from the fertilization envelope by the synthesis and secretion of hatching enzymes and starts swimming freely by ciliary activity (Stokes and Holland, 1995; Stokes, 1997). The neural plate is V-shaped (**Supplementary Figure 1E**) and the primordium of the notochord is a round mass of cells extending ventrally along the neural plate (**Supplementary Figure 1E′**). Central nervous system, notochord, and somites are clearly distinguishable, although the boundaries between notochord and somites are not always evident (**Figure 3C′** and **Supplementary Figures 1E,E′,E′′**; Hirakow and Kajita, 1994). The archenteron located anterior to the first somite pair starts expanding at this stage, forming two dorsolateral lobes (**Supplementary Figure 1E′′**).

At the N3 stage, the embryo is characterized by 6–7 somite pairs (**Figures 3D,D′**). The neural tube is closing, but will only become circular at subsequent developmental stages. The notochord is individualized from the somites, except at the most anterior tip of the embryo (Hatschek, 1893; Conklin, 1932). Ventral extensions of the somites start to generate the lateral and ventral coeloms as well as the musculature of the atrial floor (Holland and Onai, 2012). Furthermore, expression of early markers of Hatschek's nephridium, such as *pax2/5/8*, becomes detectable in the mesothelial wall of the first somite on the left side of the embryo (Kozmik et al., 1999, 2007; Carvalho et al., 2017a). The subsequent N4 stage is characterized by 8–9 somite pairs (**Figures 3E,E′**). At this stage, the two dorsolateral lobes that originated from the anterior archenteron have formed two distinctive head cavities: Hatschek's left and right diverticulum (Willey, 1894; Grassé, 1948).

The N5 stage, which is characterized by 10–11 somite pairs, is when the asymmetric formation of somites from the tail bud is initiated (**Figures 3F,F′**). Thus, while early somites are established by enterocoely from endomesoderm internalized during gastrulation, starting at the N5 stage, somites are formed by schizocoely from the tail bud (**Supplementary Figure 2**; Holland, 2015). At this stage, Hatschek's left and right diverticulum are asymmetrically organized: while the left

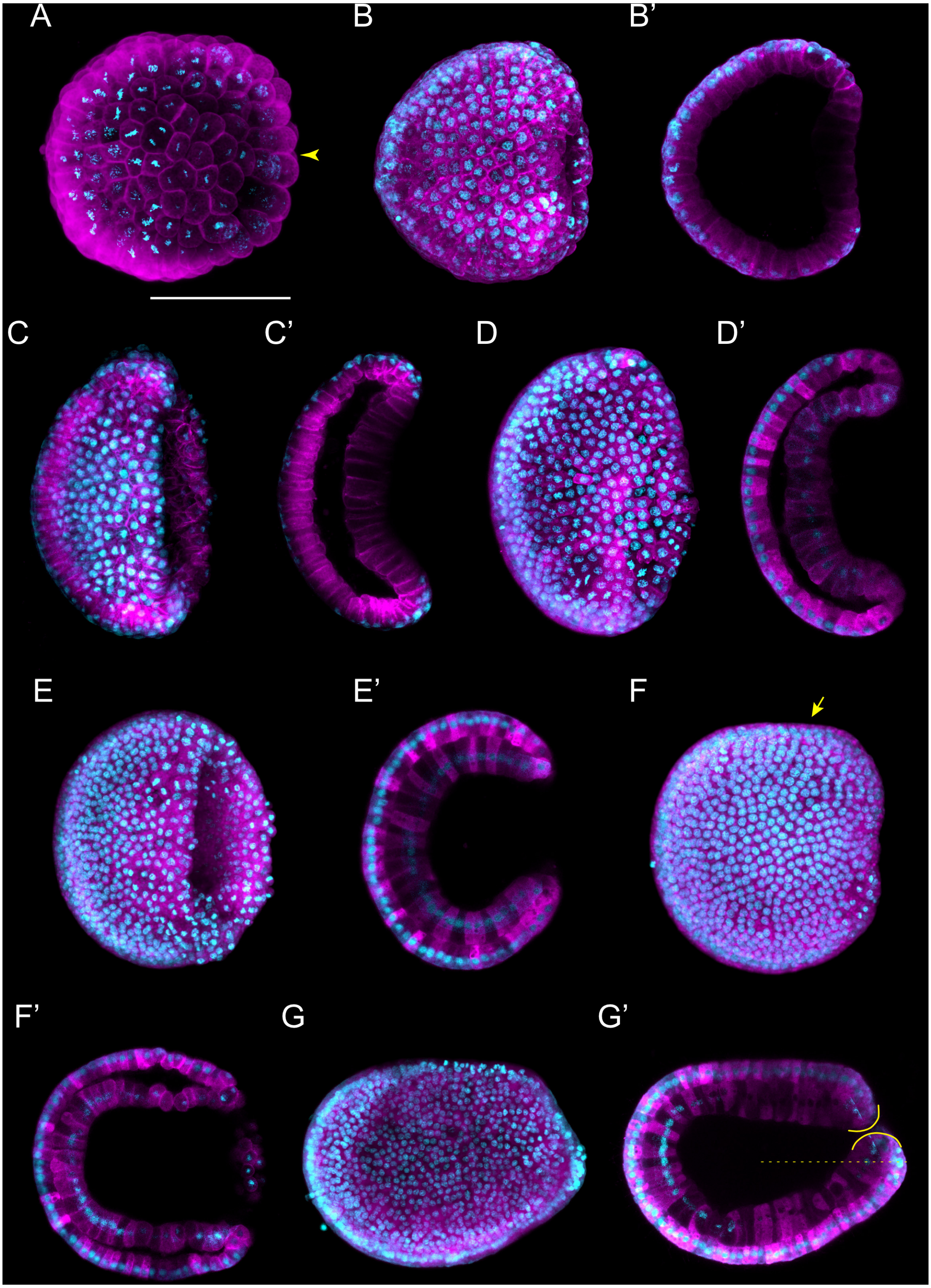


FIGURE 2 | *Branchiostoma lanceolatum* gastrula stages. Embryos are stained with the lipophilic dye FM 4-64 (magenta) and with the DNA dye Hoechst (cyan). Animal pole and anterior are to the left and dorsal side is up. **(A–G)** Maximum projections of confocal z-stacks of entire embryos. **(B′–G′)** Single z-stacks highlighting the inner morphology of the developing gastrula. **(A)** G0 stage, **(B,B′)** G1 stage, **(C,C′)** G2 stage, **(D,D′)** G3 stage, **(E,E′)** G4 stage, **(F,F′)** G5 stage, **(G,G′)** G6 stage. In **(A)**, the yellow arrowhead indicates the vegetal cells. In **(F)**, the yellow arrow highlights the flattened side of the gastrula embryo. In **(G′)**, the yellow lines delimit the upper and lower lips of the blastopore, and the dashed line indicates the midline of the embryo. Scale bar: 50 μm.

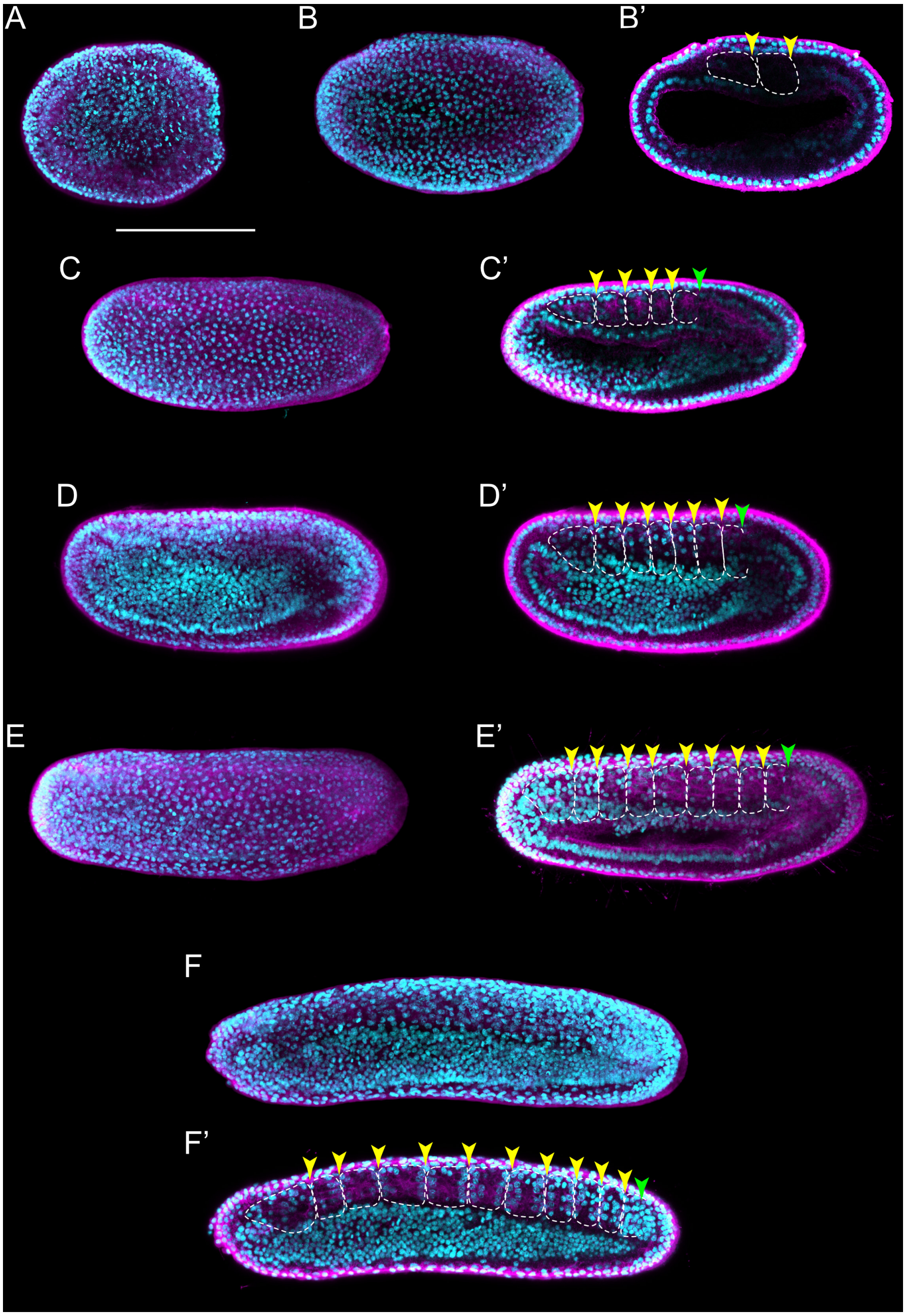


FIGURE 3 | *Branchiostoma lanceolatum* neurula stages. Embryos are labeled for aPKC (magenta) and stained with the DNA dye Hoechst (cyan). Anterior pole is to the left and dorsal side is up. **(A–F)** Average projections for aPKC (magenta) and maximum projections for Hoechst DNA staining (cyan) of confocal z-stacks of entire embryos. **(B′–F′)** Single z-stacks highlighting the inner morphology of the developing neurula. **(A)** N0 stage, **(B,B′)** N1 stage, **(C,C′)** N2 stage, **(D,D′)** N3 stage, **(E,E′)** N4 stage, **(F,F′)** N5 stage. In **(B′–F′)**, white dashed lines delineate the somites, the yellow arrowheads indicate the posterior limit of the somites and the green arrowheads highlight the posterior limit of somites newly established by enterocoely **(C′–E′)** or newly formed by schizocoely **(F′)**. Scale bar: 100 μm.

diverticulum roughly maintains its original form and size, the right diverticulum moves anteriorly, flattens, and increases in size (Willey, 1894). Furthermore, the primordium of the club-shaped gland is first discernable, ventrally in the anterior endoderm on the right side of the embryo. This developmental stage is further characterized by a decrease of proliferative activity in somites and notochord, where it becomes limited to cells at the posterior end of the embryo. However, cell proliferation continues in the tail bud, in the endoderm, and in the anterior neural plate (Holland and Holland, 2006).

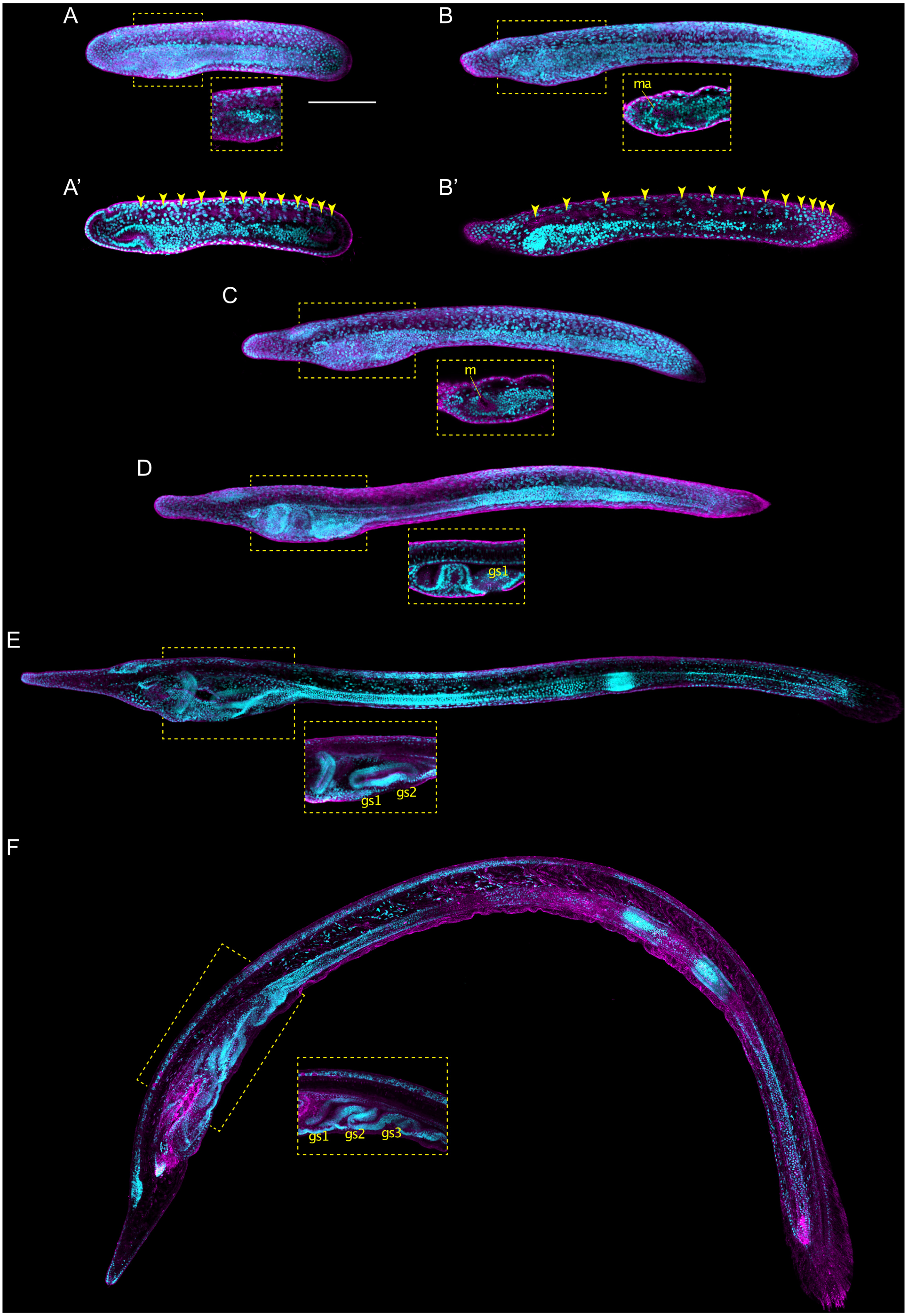


FIGURE 4 | *Branchiostoma lanceolatum* tailbud and larval stages. Embryos and larvae are labeled for aPKC (magenta) and stained with the DNA dye Hoechst (cyan). Average projections for aPKC (magenta) and maximum projections for Hoechst DNA staining (cyan) of confocal z-stacks of entire embryos and larvae. Anterior pole is to the left and dorsal side is up. **(A,A′)** T0 stage, **(B,B′)** T1 stage, **(C)** L0 stage, **(D)** L1 stage, **(E)** L2 stage, **(F)** L3 stage. **(A′,B′)** Single z-stacks highlighting the inner morphology of the developing embryo. Insets in **(A–F)** highlight the pharyngeal region. In **(A′,B′)**, yellow arrowheads indicate the posterior limits of the somites. gs1, 1st gill slit; gs2, 2nd gill slit; gs3, 3rd gill slit; m, mouth; ma, mouth anlagen. Scale bar: 100 µm.

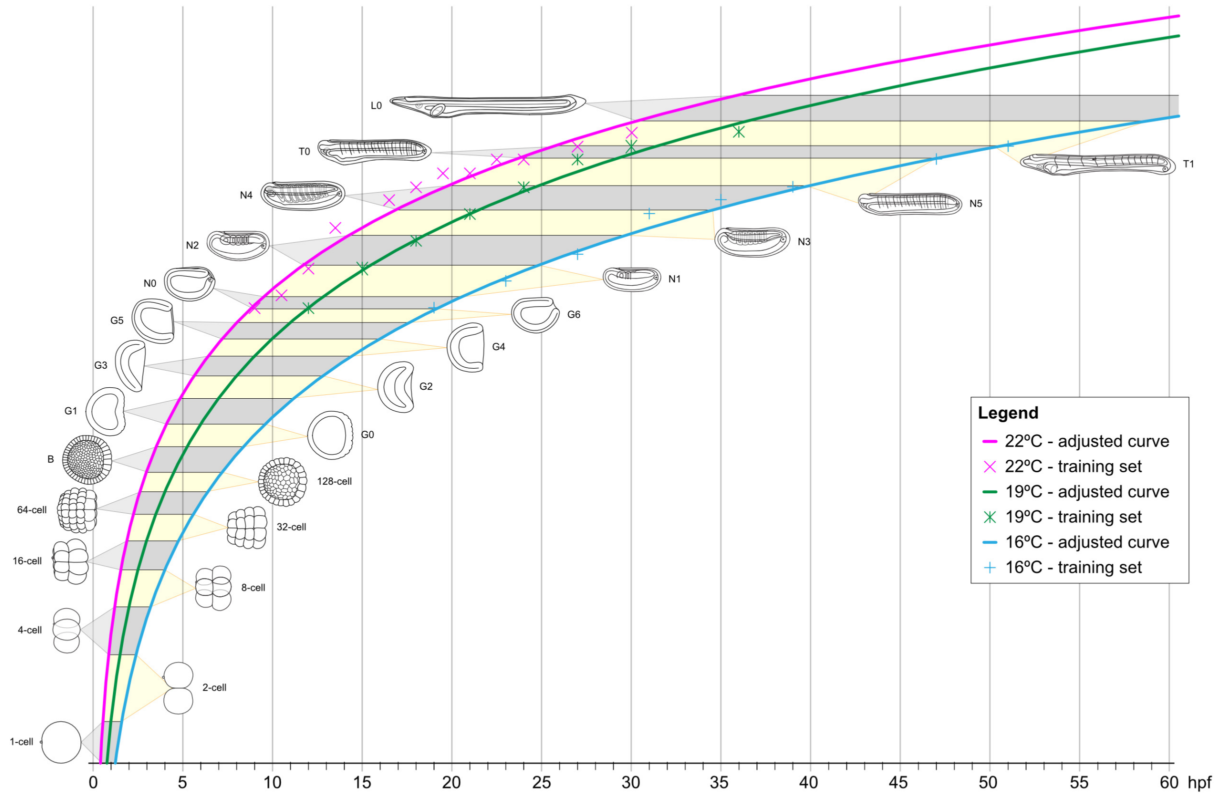


FIGURE 5 | Growth curves of *Branchiostoma lanceolatum* embryos and larvae at 16°C, 19°C, and 22°C. Animal schematics and stage nomenclatures are according to the staging system detailed in **Figure 6**. Tendency adjusted curves were obtained from the training sets and are defined by the equations: [y = 12.403ln(x) - 36.493] for 16°C; [y = 12.466ln(x) - 30.812] for 19°C; [y = 11.25ln(x) - 24.354] for 22°C. These curves use natural logarithms and do thus not reach 0 h post fertilization (0 hpf). The graphs were simplified accordingly. hpf, hours post fertilization.

Tailbud and Larva

Following neurulation, at the T0 stage, the embryo has 12 pairs of somites and exhibits a transitional morphology between neurula and larva stages (**Figures 4A,A′**) that resembles a generic vertebrate tailbud stage embryo (Slack et al., 1993; Marlétaz et al., 2018). At this T0 stage, the anterior portion of the embryo becomes clearly distinct from the posterior one, as the pharyngeal region commences to grow. In addition, the embryo starts to twitch and bend as its neuromuscular system slowly becomes operational (Hirakow and Kajita, 1994; Zieger et al., 2017). At the subsequent T1 stage, embryos are longer than those at the T0 stage, but this length difference is not due to the addition of a significant number of new somite pairs. Instead, it is due to the maturation and elongation of the existing ones, in particular those located in the anterior half of the embryo (**Figures 4B,B′** and **Supplementary Figures 3A,B**). Thus, while the mean length of a somite in somite pairs 2–7 at the T0 stage is about 30.27 µm, it increases to about 51.65 µm at the T1 stage. The overall shape of the embryo also changes at the T1 stage: the body is becoming slender as the embryo elongates, a distinctive rostral snout is appearing and the tail fin is starting to form in the caudal ectoderm (**Supplementary Figure 3B**; Hirakow and Kajita, 1994). The first pigment spot in the central nervous system appears, located in the ventral wall of the neural tube at the level of the fifth somite pair (**Supplementary Figure 3B**; Willey, 1894). Concomitant with the elongation of the rostral snout, the right diverticulum expands anteriorly, hence forming the snout cavity below the notochord (**Supplementary Figures 3A,C**). In addition, the left diverticulum starts fusing with the ectoderm to form the pre-oral pit, and the anlage of the mouth is clearly visible. Yet, neither one of these two structures penetrates the ectoderm and opens to the exterior at this stage (Kaji et al., 2016).

The earliest larva, the L0 stage, already features the main structural elements that define the asymmetry, along the left-right axis, of all subsequent larval stages (**Figure 4C**). The larval mouth opens on the left side of the developing animal by fusion of ectoderm and endoderm (**Figures 4C**; Kaji et al., 2016; Holland, 2018). The left diverticulum has now penetrated the ectoderm to form the pre-oral pit, also known as Hatschek's pit (**Supplementary Figure 3C**). Hatschek's nephridium, the kidney of larval lancelets, is now detectable between the ectoderm and the anterior-most somite on the left side of the larva (Hatschek, 1893; Holland, 2018). On the right side, the club-shaped gland is forming in the anterior endoderm, opposite to the mouth (**Supplementary Figure 3C**; Goodrich, 1930). Once completely developed, the club-shaped gland resembles a tube that connects the pharyngeal lumen on the right with the external environment

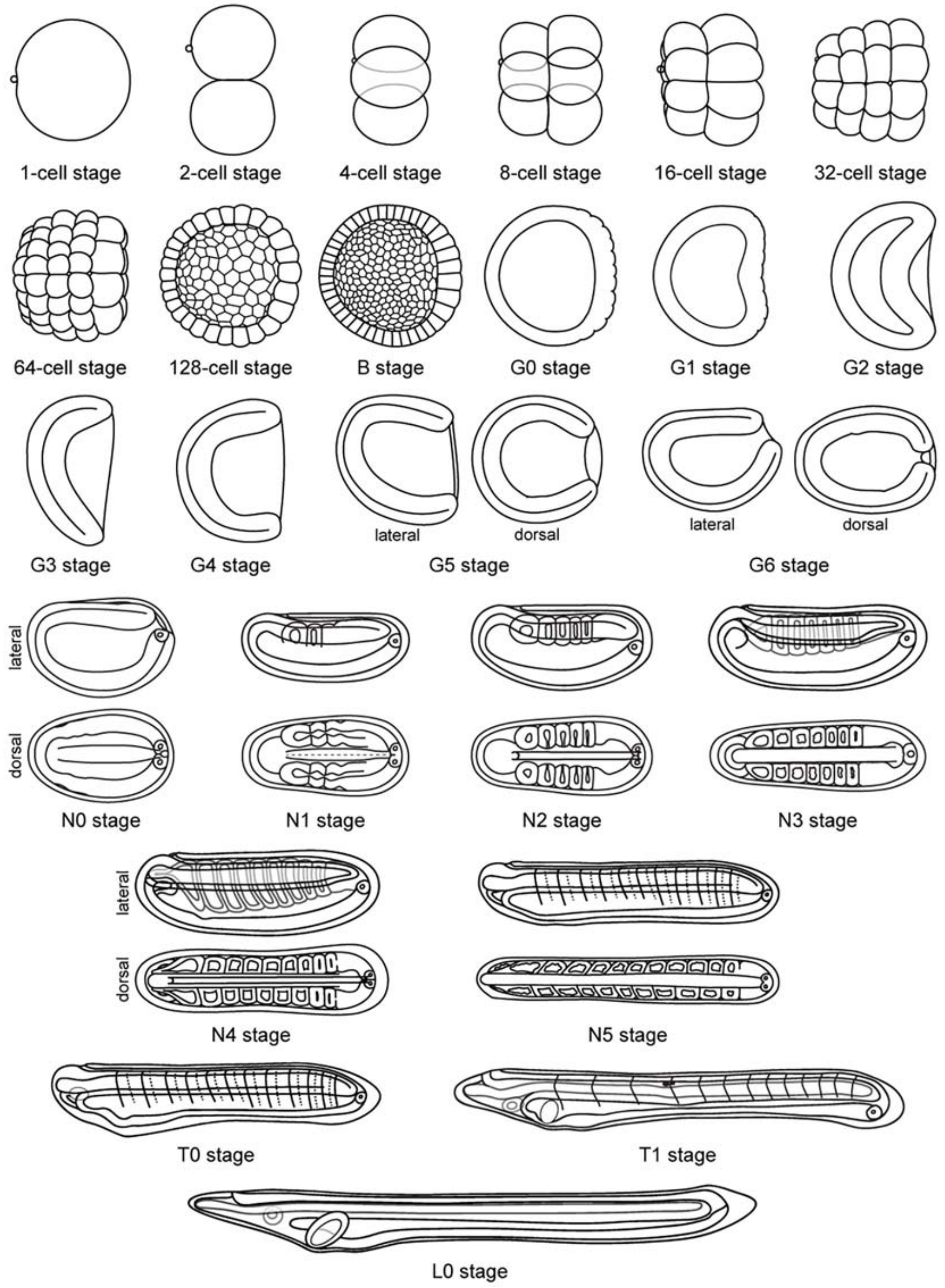

1-cell stage
2-cell stage
4-cell stage
8-cell stage
16-cell stage
32-cell stage
64-cell stage
128-cell stage
B stage
G0 stage
G1 stage
G2 stage
G3 stage
G4 stage
lateral
dorsal
G5 stage
lateral
dorsal
G6 stage
lateral
dorsal
N0 stage
N1 stage
N2 stage
N3 stage
lateral
dorsal
N4 stage
N5 stage
T0 stage
T1 stage
L0 stage

FIGURE 6 | Schematic representation of *Branchiostoma lanceolatum* development. Representations from the 1-cell stage to the L0 stage. Animal pole and anterior pole are to the left and dorsal side is up in lateral views. Drawings adapted from Hatscheck's original descriptions of *B. lanceolatum* development (Hatschek, 1881).

on the left (Jefferies, 1987). The opening is located just anterior to the mouth and is characterized by cells bearing large cilia that create a water current from the exterior into the organ (Olsson, 1983). The club-shaped gland has been shown to secrete mucoproteins and might thus contribute to larval feeding (Holland, 2015). Another structure detectable on the right side of the pharynx at the L0 stage is the endostyle. The endostyle forms from a thickening of the endodermal wall and is located just anterior to the club-shaped gland (**Supplementary Figure 3C**). The endostyle, which secretes mucous used to trap food particles, has been proposed to be homologous to the vertebrate thyroid gland (Ogasawara, 2000; Paris et al., 2008; Bertrand and Escrivá, 2011).

Although the definitive gill slits of lancelet larvae are found on the right side of the body (Holland, 2015), the anlage of the first gill slit forms at the ventral midline

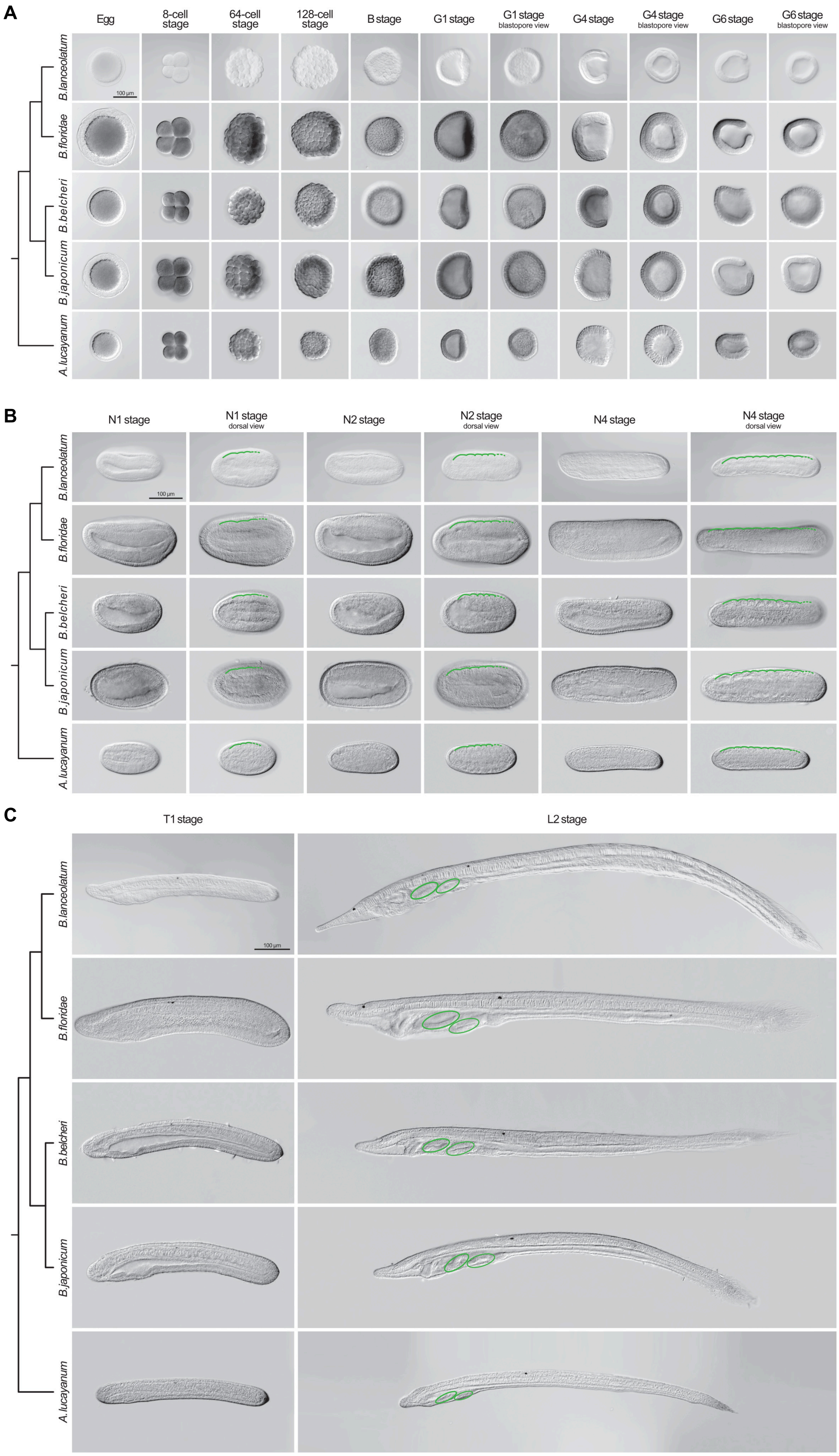

A
Egg
8-cell stage
64-cell stage
128-cell stage
B stage
G1 stage
G1 stage blastopore view
G4 stage
G4 stage blastopore view
G6 stage
G6 stage blastopore view
B.lanceolatum
B.floridae
B.belcheri
B.japonicum
A.lucayanum
100 µm
B
N1 stage
N1 stage dorsal view
N2 stage
N2 stage dorsal view
N4 stage
N4 stage dorsal view
B.lanceolatum
B.floridae
B.belcheri
B.japonicum
A.lucayanum
100 µm
C
T1 stage
L2 stage
B.lanceolatum
B.floridae
B.belcheri
B.japonicum
A.lucayanum
100 µm

FIGURE 7 | Comparison of lancelet development. Five species were analyzed: *Branchiostoma lanceolatum*, *Branchiostoma floridae*, *Branchiostoma belcheri*, *Branchiostoma japonicum*, and *Asymmetron lucayanum*. **(A)** cleavage, blastula, and gastrula stages, **(B)** neurula stages, **(C)** tailbud and larva stages. Cladograms represent the evolutionary relationship between the different species (Igawa et al., 2017). The green lines in **(B)** trace the somites on one side of the neurula, with dashed green lines highlighting forming somites. The green ovals in **(C)** indicate the gill slits of the larva. Scale bars: 100 µm.

TABLE 1 | Comparison of lancelet development.

		B. lanceolatum ^{1}			***B. floridae*** ^{2}		***B. belcheri*** ^{3}	***B. japonicum*** ^{4}	***A. Lucayanum*** ^{5}
Stage	**Key feature**	***at 16°C***	***at 19°C***	***at 22°C***	***at 25°C***	***at 30°C***	***at 24°C***	***at 25°C***	***at 27°C***
1-cell stage	fertilized egg								
2-cell stage	2 cells		1 h		45′	30′	45′	50′-1 h	1 h-1 h30′
4-cell stage	4 cells		1 h30′		1 h	50′	1 h	1 h10′	2 h-2 h30′
8-cell stage	8 cells		2 h		1 h30′	1 h	1 h20′	1 h35′	2 h-2 h30′
16-cell stage	16 cells		2 h30′		2 h	1 h15′	1 h45′	1 h55′	3 h-3 h30′
32-cell stage	32 cells		3 h		2 h15′	1 h30′	2 h10′	2 h15′	3 h-3 h30′
64-cell stage	64 cells		3 h30′		2 h30′	1 h45′	2 h30′	2 h35′	4 h-4 h30′
128-cell stage	128 cells		4 h		3 h	2 h	2 h50′	3 h10′	4 h-4 h30′
B stage	initiation of asynchronous cell division		4 h30′		4 h	2 h30′		3 h20′	5 h
G0 stage	initial flattening of the vegetal zone		5 h				3 h30′		
G1 stage	flattened vegetal pole		6 h		4 h30′	3 h30′		3 h40′	
G2 stage	invaginated vegetal pole		7 h					4 h10′	
G3 stage	cap shaped		8 h		5 h	4 h	3 h55′	5 h35′	
G4 stage	cup shaped	15 h	9 h		6 h	4 h30′		6 h-6 h20′	9 h
G5 stage	vase shaped		10 h				5 h45′	7 h40′	
G6 stage	bottle shaped		11 h		6 h30′	5 h	7 h55′	8 h50′	
N0 stage	no somite pairs, neural plate	19 h	12 h	9 h			8 h30′		12 h
N1 stage	1–3 somite pairs	23 h	15 h	12 h	8 h30′	6 h	10 h20′	10 h30′	15 h
N2 stage	4–5 somite pairs, hatching	27 h	18 h		9 h30′	6 h30′	12 h	13 h	19 h
N3 stage	6–7 somite pairs	31 h	21 h	13 h30′					
N4 stage	8–9 somite pairs, prior to schizocoelic somite formation	35 h	24 h	16 h30′	10 h30′	7 h30′		18 h	
N5 stage	10–11 somite pairs	47 h	27 h	19 h30′					32 h
T0 stage	12 somite pairs, tailbud shape, enlarged pharyngeal region	51 h	30 h	27 h					
T1 stage	13 somite pairs, mouth and pre-oral pit anlagen, first pigment spot		36 h	30 h	20 h/24 h	12 h/17 h		24 h	50 h
L0 stage	no gill slits, open mouth		42 h		30 h	21 h			
L1 stage	1 gill slit		48 h		32 h/36 h	23 h/28 h		36 h	72 h
L2 stage	2 gill slits				42 h/72 h	36 h	36 h/48 h	48 h	
Ln stage	n gill slits								

Species: Branchiostoma lanceolatum, Branchiostoma floridae, Branchiostoma belcheri, Branchiostoma japonicum, Asymmetron lucayanum. Data origin: ^{1}*Current study,* ^{2}*Stokes and Holland, 1995; Holland and Holland, 1998; Holland and Yu, 2004; Holland et al., 2015,* ^{3}*Zhang, 2017,* ^{4}*Hirakow and Kajita, 1990, 1991, 1994; Morov et al., 2016,* ^{5}*Holland and Holland, 2010; Holland et al., 2015. "/" indicates that different developmental times have been reported.*

at the L0 stage (**Supplementary Figure 3C**). The anlage of the anus arises at the same stage at the posterior end of the gut, which is located just anterior to the ectodermal caudal fin (**Supplementary Figure 3C**; Jefferies, 1987). However, while the anlage of the anus also originates at the ventral midline, the definitive anus will be located on the left side of the body (Jefferies, 1987). The first definitive gill slit penetrates at the L1 stage (**Figure 4D**), and, following the establishment of all the structures referred to above, the L1 larva starts feeding. Following the L1 stage, new gill slits are added sequentially, hence defining the subsequent developmental stages: L2 stage for 2 gill slits (**Figure 4E**), L3 stage for 3 gill slits (**Figure 4F**), and so on, until the larva enters metamorphosis. The number of gill slits required before a larva becomes competent to undergo metamorphosis varies between different lancelet species (Wickstead, 1967; Holland and Yu, 2004; Urata et al., 2007; Carvalho et al., 2017b).

Branchiostoma lanceolatum Developmental Timing

It is well established that temperature directly affects the speed and potentially even the progression of animal development, in lancelets as well as in other animals (Fuentes et al., 2007; Ebisuya and Briscoe, 2018). To define the impact of temperature on *B. lanceolatum* development, we reared embryos and larvae at three different temperatures (16°C, 19°C, and 22°C). We then mapped their developmental progression, according to our staging system and using somite pairs as defining landmark. To visualize the somites, embryos were fixed at regular intervals starting at the N0 stage, and *in situ* hybridization was performed with the somite marker *mrf1*. For each of the three temperatures, the number of somite pairs at a given developmental time was subsequently used as a training set (**Supplementary Figure 4** and **Supplementary Table 1**) to define the growth curve that best reflected *B. lanceolatum* development. We further extrapolated the time intervals for the different development stages of our staging system prior to and following the neurula stages (**Figure 5**). The results show that, despite a marked effect on the speed of development, the shapes of the growth curves, marking the progression of development, are very similar for the three temperatures (**Figure 5**). This indicates that the different temperatures predominantly impact the rate of cell division during development and not the overall physiology of the embryos and larvae. It is, however, almost certain that *B. lanceolatum* can only develop within a certain temperature range. *B. lanceolatum* adults, for example, die after being cultured at 30°C for 2 weeks (Fuentes et al., 2007), and it is likely that embryos and larvae are even more temperature sensitive than adults. The results further demonstrate that these growth curves can be used to easily transform a developmental stage expressed as time after fertilization into an unambiguous stage name.

Comparative Lancelet Developmental Staging

We next assessed whether the staging table we elaborated using *B. lanceolatum* (**Figure 6**) can be applied to the development of other lancelets. For this, we compared *B. lanceolatum* embryos and larvae with those from four additional lancelet species, three from the genus *Branchiostoma* (*B. floridae*, *B. belcheri*, and *B. japonicum*) and one from the genus *Asymmetron* (*A. lucayanum*). A total of 13 developmental stages were included in the comparative analysis: unfertilized eggs, 8-cell, 64-cell, 128-cell, B, G1, G4, G6, N1, N2, N4, T1, and L2 (**Figure 7**). DIC images of the different stages revealed a strong overall conservation of the morphology of the five species. However, differences were detected in the overall size of the developing lancelets. The unfertilized egg of *B. floridae*, for example, is significantly larger than those of the other analyzed species. The diameter of the *B. floridae* egg is about 25% larger than that of *B. lanceolatum*, 18% larger than that of *B. belcheri*, 22% larger than that of *B. japonicum*, and 33% larger than that of *A. lucayanum* (**Figure 7A**). Another notable difference is the appearance of pigmentation in the posterior-most ectoderm, which is detectable as early as the N4 stage in *A. lucayanum*, but only appears at the T1 stage in the *Branchiostoma* species (**Figures 7B,C**; Zieger et al., 2017). In addition, the timing of rostrum and tail fin formation is not strictly conserved (**Figure 7C**). Thus, while the rostrum is clearly elongated in T1 stage *B. lanceolatum*, development of the snout region is much less advanced in the other species, in particular in *A. lucayanum* (**Figure 7C**). The lack of anterior head cavities in members of the genus *Asymmetron* may at least partially explain this prominent difference (Holland and Holland, 2010; Holland et al., 2015). Posteriorly, pigmented cells are detectable in *A. lucayanum* as well as *B. lanceolatum* and *B. belcheri*. In these three lancelet species, the rudiment of the forming tail fin is also already present at the T1 stage (**Figure 7C**). In the larva, the species-specific differences in the snout and tail regions become even more accentuated. While *B. lanceolatum* larvae have a particularly long and thin snout, the rostrum of the other lancelet species is much less pronounced. At the L2 stage, the tail fins are either pointy (in *A. lucayanum*, *B. lanceolatum*, and *B. belcheri*) or roundish (in *B. floridae* and *B. japonicum*). Previous studies have further shown that, when compared to *B. floridae*, *B. lanceolatum* larvae are characterized by a heterochronic delay of second gill slit formation and that this delay is not due to differences in developmental speed (Somorjai et al., 2008).

Despite these differences, the defining characters of each developmental stage that we established in *B. lanceolatum* embryos and larva were conserved in all other lancelet species. The cleavage, gastrula, and neurula stages of the five lancelet species are thus remarkably similar (**Figures 7A,B**). Furthermore, the rate of somite formation as well as the timing of appearance of key morphological features at the neurula and tailbud stages are comparable (**Figures 7B,C**). For example, the N2 stage embryo of all five species is characterized by 4–5 somite pairs, a neuropore, and a neurenteric canal. Taken together, although there are minor species-specific differences, the overall development of the five lancelets is highly conserved and fully compatible with our updated staging and stage nomenclature systems. We thus expect these systems to be widely applicable to embryos and larvae of all extant lancelets.

DISCUSSION

In the present study, we carried out a detailed analysis of the development of the lancelet *B. lanceolatum* using confocal microscopy and we defined straightforward staging and nomenclature systems for developing lancelets. We validated the updated staging system at different rearing temperatures for *B. lanceolatum* and demonstrated that it can be used for staging lancelets from the genus *Branchiostoma* as well as from the genus *Asymmetron*. This work thus resolves two fundamental problems for studies carried out in lancelets: (1) the lack of comparability between embryos and larvae from different species and (2) the confusion created by varying staging and stage nomenclature systems in a given species. Importantly, the morphological characters used to define each stage are generally easy to identify, such as the total number of cells for the cleavage stages, the initiation of asynchronous cell division for the blastula (B) stage, the shape of the gastrula (G), the number of somite pairs in the neurula (N) and tailbud (T) stages, and the formation of pharyngeal structures for the tailbud (T) and larva (L) stages. Most of these characters have previously been validated as distinguishing hallmarks of lancelet development (Kovalevsky, 1867; Hatschek, 1893; Cerfontaine, 1906; Conklin, 1932; Hirakow and Kajita, 1990, 1991, 1994) and are also regularly used for the staging of other model organisms (Kimmel et al., 1995; Richardson and Wright, 2003).

Our updated staging system also allowed us to clarify previously unresolved controversies about lancelet development. One example is the definition of the blastula stage. Some authors suggested that the blastula is established as soon as the blastocoel is enclosed by cells (at the 64-cell stage) (Holland and Yu, 2004), while others proposed that the blastula forms after the 8th round of cell divisions (after the 128-cell stage) (Hirakow and Kajita, 1990). Here, we redefined the B stage, which is characterized by the initiation of asynchronous cell divisions (at the transition from 128 cells to 256 cells) and ends with the initial flattening of the vegetal side of the embryo. In chordates, the first asynchronous cell divisions are often observed around the mid-blastula transition (MBT) and are thus correlated with the activation of zygotic gene transcription (McDougall et al., 2019). A detailed analysis of transcriptomes obtained at different developmental stages suggests that this is also the case in lancelets, as the transition from 128 cells to 256 cells is marked by a strong increase in the expression of genes required for the initiation of zygotic transcription, including, for example, those encoding nuclear ribonucleic proteins (Yang et al., 2016).

Another ambiguous developmental period is the transition between the gastrula and the neurula stage, sometimes referred to as a very late gastrula (Hirakow and Kajita, 1991) or a very early neurula (Lu et al., 2012; Zhang et al., 2013). We redefined this important stage as N0, corresponding to an embryo with a small blastopore, which is characteristic for gastrula stages, and a flattened neural plate, marking the onset of neurulation. We further expanded the classification of neurulae to six independent N stages, hence allowing more detailed descriptions of the morphological changes occurring during this crucial developmental period. Previous descriptions distinguished only three (Hirakow and Kajita, 1994) or four different N stages (Lu et al., 2012).

Another controversial point of lancelet development is the definition of the larva. Some authors claimed that the larval stage starts when "tissues and cells prepare for performing their own function" (Hirakow and Kajita, 1994). Alternatively, the larval stage has been defined by the opening of the mouth and thus by the moment the animal starts feeding (Holland, 2015). To clarify this issue, we defined a new developmental period for lancelets that, based on the gestalt of the embryo at this stage, we called the tailbud (T) stage (Lemaire, 2011). We further defined the onset of the larval stage (L0) as the moment when the mouth opens, as it has previously been suggested for lancelets (Holland, 2015) and other animals (Kimmel et al., 1995; Smith et al., 2008).

Significant efforts have been made to develop protocols for maintaining and spawning adult lancelets in captivity and for manipulating lancelet embryos and larvae. Thanks to these efforts, lancelets have become attractive laboratory models (Li et al., 2013, 2014; Carvalho et al., 2017b; Shi et al., 2018; Su et al., 2020). However, one of the remaining obstacles was the absence of a widely applicable staging system guaranteeing the comparability of results obtained in different lancelet species. Here, we propose a complete staging system for developing lancelets. Although the stage descriptions were carried out in *B. lanceolatum*, our comparisons with other lancelet species clearly demonstrate that both staging and nomenclature are valid beyond *B. lanceolatum* and are likely applicable to all extant lancelets. Using the defining characters for each stage, we were thus able to establish a comparative developmental table for the five lancelet species used in this study: *B. lanceolatum*, *B. floridae*, *B. belcheri*, *B. japonicum*, and *A. lucayanum* (**Table 1**). In this regard, this work adds morphological evidence to genomic results suggesting that lancelets evolve at a very slow rate (Putnam et al., 2008; Igawa et al., 2017; Marlétaz et al., 2018; Simakov et al., 2020). Taken together, we strongly believe that this description and organization of embryonic and larval development, along with the ontology for anatomy and development for the *Branchiostoma* genus (AMPHX) (Bertrand et al., 2021), should become the standards for the scientific community in an effort to harmonize research on developing lancelets. We also anticipate that this updated description of lancelet development will facilitate future comparative studies between lancelets and other chordates.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for this study because lancelets are invertebrate chordates.

AUTHOR CONTRIBUTIONS

JEC and MS designed the study, conceived the experiments, analyzed the data, and wrote and revised the manuscript. JEC and FL performed the experiments in *B. lanceolatum*. LWY and J-KY performed the experiments in *B. floridae*, *B. belcheri*, *B. japonicum*, and *A. lucayanum*. JEC compiled the figures. JCC, HE, and J-KY assisted in the writing and editing of the manuscript. All authors have read and approved the final version of the manuscript.

FUNDING

JEC was a FCT doctoral fellow (SFRH/BD/86878/2012) and is currently supported by a FRM fellowship (SPF20170938703). LWY and J-KY are supported by Academia Sinica intramural funds and grants from the Ministry of Science and Technology, Taiwan (MOST-105-2628-B-001-003-MY3 and MOST-108-

2311-B-001-035-MY3). HE is financed by ANR-16-CE12-0008-01, ANR-17-CE13-0027-02, and ANR-19-CE13-0011. FL, JCC, and MS are funded by the CNRS.

ACKNOWLEDGMENTS

We are indebted to Linda Z. Holland and Nicholas D. Holland from the Scripps Institution of Oceanography, La Jolla, CA, United States, for collecting *A. lucayanum* adults. Janet Chenevert from the Laboratoire de Biologie du Développement de Villefranche-sur-Mer, Villefranche-sur-Mer, France, kindly provided the FM 4-64 lipophilic dye and the primary antibody against aPKC as well as useful technical advice. We would like to thank Estelle Hirsinger from the Institut de Biologie Paris-Seine, Paris, France, for reading and commenting the manuscript and Tzu-Kai Huang and the staff at the Marine Research Station of the Institute of Cellular and Organismic Biology, Taipei, Taiwan, for technical assistance. We would also like to thank the Centre de Ressources Biologiques (CRB) of the Institut de la Mer de Villefranche (IMEV), Villefranche-sur-Mer, France specifically the Service Aquariologie (SA) and the Mediterranean Culture Collection of Villefranche (MCCV), and the Plateforme d'Imagerie par Microscopie (PIM) of the Institut de la Mer de Villefranche (IMEV), Villefranche-sur-Mer, France which are supported by EMBRC-France (ANR-10-INBS-02). This article is a back-to-back submission with Bertrand et al., 2021.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.668006/full#supplementary-material

Supplementary Figure 1 | Detailed highlights of specific structures of *Branchiostoma lanceolatum* development during cleavage and neurula stages. **(A,B)** Embryos are stained with the lipophilic dye FM 4-64 (magenta). **(C–E)** Embryos are labeled for aPKC (magenta) and stained with the DNA dye Hoechst (cyan). The anterior pole is to the left, and, on the dorsal views, the right side is up **(C,C′,E–E′′)**, while, on the lateral view, the dorsal side is up **(D)**. Maximum projections of confocal z-stacks of *B. lanceolatum* embryos at the 1 cell-stage **(A)**, 32-cell stage **(B)**, N1 stage **(C,C′)**, and N2 stage **(D–E′′)**. Insets **(a–d)** in **(A)** correspond to regions highlighted with dotted rectangles and are shown at 2x magnification. bc, blastocoel; bp, blastopore; cv, cerebral vesicle; m, maternal DNA; nc, neurenteric canal; np, neuropore; nrt, neural tube; nt, notochord; p, paternal DNA; pb1, 1st polar body; pb2, 2nd polar body; phc, presumptive head cavities; s1–5, somite pairs 1 to 5; sm, somitic mesoderm. Scale bar: 50 μm.

Supplementary Figure 2 | Detailed highlights of specific structures of *Branchiostoma lanceolatum* development during neurula stages. **(A,B)** Embryos are labeled for aPKC (magenta) and stained with the DNA dye Hoechst (cyan), with the colors having been inverted in the insets for clarity. Insets correspond to regions highlighted with dotted rectangles and are shown at 2x magnification. The anterior pole is to the left, and the dorsal side is up. Single z-stacks show the formation of a new somite by enterocoely at the N4 stage **(A)** and by schizocoely from the tail bud at the N5 stage **(B)**. Scale bar: 50 μm.

Supplementary Figure 3 | Detailed highlights of specific structures of *Branchiostoma lanceolatum* development during tailbud and larval stages. Embryos and larvae are labeled for aPKC (magenta) and stained with the DNA dye Hoechst (cyan). Embryos and larvae are in lateral views, the anterior pole is to the left and the dorsal side is up. T0 **(A)**, T1 **(B)**, and L0 **(C)** stages are shown. Insets **(a–g)** in **(C)** correspond to regions highlighted with dotted rectangles and are shown at 2x magnification. an, anus; cc, cephalic coelom; csg, club-shaped gland; cv, cerebral vesicle; en, endostyle; np, neuropore; nrt, neural tube; nt, notochord; pgs, presumptive 1st gill slit; pp, pre-oral pit; ps, 1st pigment spot; rd, right diverticulum; s2–5, somite pairs 2 to 5; tf, tail fin. Scale bar: 50 μm.

Supplementary Figure 4 | Expression of the *mrf1* gene in developing *Branchiostoma lanceolatum* reared at different temperatures. Embryos are in dorsal views with anterior pole to the left and right side up. **(A)** 16°C, **(B)** 19°C, and **(C)** 21°C. On each image, the time of development in hours post fertilization (h) and the number of fully formed somite pairs (s) are indicated. Scale bars: 50 μm.

Supplementary Table 1 | Somite pair counts based on the expression of the *mrf1* gene in developing *Branchiostoma lanceolatum* reared at three different temperatures (16°C, 19°C, and 21°C) (**Supplementary Figure 4**), and natural logarithmic tendency curves obtained from the three training sets.

REFERENCES

Aase-Remedios, M. E., Coll-Lladó, C., and Ferrier, D. E. K. (2020). More than one-to-four via 2R: evidence of an independent amphioxus expansion and two-gene ancestral vertebrate state for MyoD-related myogenic regulatory factors (MRFs). *Mol. Biol. Evol.* 37, 2966–2982. doi: 10.1093/molbev/msaa147

Acemel, R. D., Tena, J. J., Irastorza-Azcarate, I., Marlétaz, F., Gómez-Marín, C., de la Calle-Mustienes, E., et al. (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat. Genet.* 48, 336–341. doi: 10.1038/ng.3497

Annona, G., Caccavale, F., Pascual-Anaya, J., Kuratani, S., De Luca, P., Palumbo, A., et al. (2017). Nitric oxide regulates mouth development in amphioxus. *Sci. Rep.* 7:8432. doi: 10.1038/s41598-017-08157-w

Barresi, M. J. F., and Gilbert, S. F. (2019). *Developmental Biology*. 12th Edn. New York, NY: Oxford University Press.

Bertrand, S., Camasses, A., Somorjai, I., Belgacem, M. R., Chabrol, O., Escande, M.-L., et al. (2011). Amphioxus FGF signaling predicts the acquisition of vertebrate morphological traits. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9160–9165. doi: 10.1073/pnas.1014235108

Bertrand, S., Carvalho, J. E., Dauga, D., Matentzoglu, N., Daric, V., Yu, J.-K., et al. (2021). The ontology of the amphioxus anatomy and life cycle (AMPHX). *Front. Cell Dev. Biol.* 9:668025. doi: 10.3389/fcell.2021.668025

Bertrand, S., and Escrivá, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Carvalho, J. E., Lahaye, F., Croce, J. C., and Schubert, M. (2017a). CYP26 function is required for the tissue-specific modulation of retinoic acid signaling during amphioxus development. *Int. J. Dev. Biol.* 61, 733–747. doi: 10.1387/ijdb.170227ms

Carvalho, J. E., Lahaye, F., and Schubert, M. (2017b). Keeping amphioxus in the laboratory: an update on available husbandry methods. *Int. J. Dev. Biol.* 61, 773–783. doi: 10.1387/ijdb.170192ms

Carvalho, J. E., Theodosiou, M., Chen, J., Chevret, P., Alvarez, S., De Lera, A. R., et al. (2017c). Lineage-specific duplication of amphioxus retinoic acid degrading enzymes (CYP26) resulted in sub-functionalization of patterning and homeostatic roles. *BMC Evol. Biol.* 17:24. doi: 10.1186/s12862-016-0863-1

Cerfontaine, P. (1906). *Recherches Sur le Développement de l'Amphioxus*. Liège: Impr. H. Vaillant-Carmanne.

Conklin, E. G. (1932). The embryology of amphioxus. *J. Morphol.* 54, 69–151. doi: 10.1002/jmor.1050540103

Ebisuya, M., and Briscoe, J. (2018). What does time mean in development? *Development* 145:dev164368. doi: 10.1242/dev.164368

Fuentes, M., Benito, E., Bertrand, S., Paris, M., Mignardot, A., Godoy, L., et al. (2007). Insights into spawning behavior and development of the European amphioxus (*Branchiostoma lanceolatum*). *J. Exp. Zool. B Mol. Dev. Evol.* 308B, 484–493. doi: 10.1002/jez.b.21179

Fuentes, M., Schubert, M., Dalfo, D., Candiani, S., Benito, E., Gardenyes, J., et al. (2004). Preliminary observations on the spawning conditions of the European amphioxus (*Branchiostoma lanceolatum*) in captivity. *J. Exp. Zool. B Mol. Dev. Evol.* 302B, 384–391. doi: 10.1002/jez.b.20025

Goodrich, E. S. (1930). Memoirs: the development of the club-shaped gland in amphioxus. *J. Cell Sci.* S2-74, 155–164.

Grassé, P.-P. (1948). *Traité de Zoologie: Anatomie, Systématique, Biologie - Tome XI, Echinodermes - Stomocordes - Procordes*. Paris: Masson & Cie.

Hatschek, B. (1881). *Studien über Entwicklung des Amphioxus*. Wien: A. Hölder.

Hatschek, B. (1893). *The Amphioxus and its Development*, ed. J. Tuckey (London: Swan Sonnenschein & Co).

Hirakow, R., and Kajita, N. (1990). An electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: cleavage. *J. Morphol.* 203, 331–344. doi: 10.1002/jmor.1052030308

Hirakow, R., and Kajita, N. (1991). Electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: the gastrula. *J. Morphol.* 207, 37–52. doi: 10.1002/jmor.1052070106

Hirakow, R., and Kajita, N. (1994). Electron microscopic study of the development of amphioxus, *Branchiostoma belcheri tsingtauense*: the neurula and larva. *Kaibogaku Zasshi* 69, 1–13.

Holland, L. Z. (2015). "*Cephalochordata*," in *Evolutionary Developmental Biology of Invertebrates*, ed. A. Wanninger (Berlin: Springer), 91–133. doi: 10.1007/978-3-7091-1856-6_3

Holland, L. Z., and Holland, N. D. (1992). Early development in the lancelet (= amphioxus) *Branchiostoma floridae* from sperm entry through pronuclear fusion: presence of vegetal pole plasm and lack of conspicuous ooplasmic segregation. *Biol. Bull.* 182, 77–96. doi: 10.2307/1542182

Holland, L. Z., and Holland, N. D. (1998). Developmental gene expression in amphioxus: new insights into the evolutionary origin of vertebrate brain regions, neural crest, and rostrocaudal segmentation. *Am. Zool.* 38, 647–658. doi: 10.1093/icb/38.4.647

Holland, L. Z., and Onai, T. (2012). Early development of cephalochordates (amphioxus). *Wiley Interdiscip. Rev. Dev. Biol.* 1, 167–183. doi: 10.1002/wdev.11

Holland, L. Z., and Yu, J.-K. (2004). Cephalochordate (amphioxus) embryos: procurement, culture, and basic methods. *Methods Cell Biol.* 74, 195–215. doi: 10.1016/s0091-679x(04)74009-1

Holland, N. D. (2018). Formation of the initial kidney and mouth opening in larval amphioxus studied with serial blockface scanning electron microscopy (SBSEM). *EvoDevo* 9:16. doi: 10.1186/s13227-018-0104-3

Holland, N. D., and Holland, L. Z. (1989). Fine structural study of the cortical reaction and formation of the egg coats in a lancelet (= amphioxus), *Branchiostoma floridae* (phylum Chordata: subphylum Cephalochordata = Acrania). *Biol. Bull.* 176, 111–122. doi: 10.2307/1541578

Holland, N. D., and Holland, L. Z. (2006). Stage- and tissue-specific patterns of cell division in embryonic and larval tissues of amphioxus during normal development. *Evol. Dev.* 8, 142–149. doi: 10.1111/j.1525-142X.2006.00085.x

Holland, N. D., and Holland, L. Z. (2010). Laboratory spawning and development of the Bahama lancelet, *Asymmetron lucayanum* (Cephalochordata): fertilization through feeding larvae. *Biol. Bull.* 219, 132–141. doi: 10.1086/BBLv219n2p132

Holland, N. D., and Holland, L. Z. (2017). The ups and downs of amphioxus biology: a history. *Int. J. Dev. Biol.* 61, 575–583. doi: 10.1387/ijdb.160395LH

Holland, N. D., Holland, L. Z., and Heimberg, A. (2015). Hybrids between the Florida amphioxus (*Branchiostoma floridae*) and the Bahamas lancelet (*Asymmetron lucayanum*): developmental morphology and chromosome counts. *Biol. Bull.* 228, 13–24. doi: 10.1086/BBLv228n1p13

Holland, N. D., Panganiban, G., Henyey, E. L., and Holland, L. Z. (1996). Sequence and developmental expression of *AmphiDll*, an amphioxus *Distal-less* gene transcribed in the ectoderm, epidermis and nervous system: insights into evolution of craniate forebrain and neural crest. *Development* 122, 2911–2920. doi: 10.1242/dev.122.9.2911

Igawa, T., Nozawa, M., Suzuki, D. G., Reimer, J. D., Morov, A. R., Wang, Y., et al. (2017). Evolutionary history of the extant amphioxus lineage with shallow-branching diversification. *Sci. Rep.* 7, 1157. doi: 10.1038/s41598-017-00786-5

Jefferies, R. P. S. (1987). *The Ancestry of the Vertebrates*. Cambridge: Cambridge University Press.

Kaji, T., Reimer, J. D., Morov, A. R., Kuratani, S., and Yasui, K. (2016). Amphioxus mouth after dorso-ventral inversion. *Zool. Lett.* 2:2. doi: 10.1186/s40851-016-0038-3

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310. doi: 10.1002/aja.1002030302

Kon, T., Nohara, M., Yamanoue, Y., Fujiwara, Y., Nishida, M., and Nishikawa, T. (2007). Phylogenetic position of a whale-fall lancelet (Cephalochordata) inferred from whole mitochondrial genome sequences. *BMC Evol. Biol.* 7:127. doi: 10.1186/1471-2148-7-127

Kovalevsky, A. O. (1867). *Entwicklungsgeschichte des Amphioxus Lanceolatus*. St Pétersbourg, XI: Mém. Académie Impérial Sci.

Kozmik, Z., Holland, N. D., Kalousova, A., Paces, J., Schubert, M., and Holland, L. Z. (1999). Characterization of an amphioxus paired box gene, *AmphiPax2/5/8*: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development* 126, 1295–1304. doi: 10.1242/dev.126.6.1295

Kozmik, Z., Holland, N. D., Kreslova, J., Oliveri, D., Schubert, M., Jonasova, K., et al. (2007). *Pax–Six–Eya–Dach* network during amphioxus development: conservation *in vitro* but context specificity *in vivo*. *Dev. Biol.* 306, 143–159. doi: 10.1016/j.ydbio.2007.03.009

Le Petillon, Y., Luxardi, G., Scerbo, P., Cibois, M., Leon, A., Subirana, L., et al. (2017). Nodal–Activin pathway is a conserved neural induction signal in chordates. *Nat. Ecol. Evol.* 1, 1192–1200. doi: 10.1038/s41559-017-0226-3

Lemaire, P. (2011). Evolutionary crossroads in developmental biology: the tunicates. *Development* 138, 2143–2152. doi: 10.1242/dev.048975

Li, G., Feng, J., Lei, Y., Wang, J., Wang, H., Shang, L. K., et al. (2014). Mutagenesis at specific genomic loci of amphioxus *Branchiostoma belcheri* using TALEN method. *J. Genet. Genom.* 41, 215–219. doi: 10.1016/j.jgg.2014.02.003

Li, G., Shu, Z., and Wang, Y. (2013). Year-round reproduction and induced spawning of Chinese amphioxus, *Branchiostoma belcheri*, in laboratory. *PLoS One* 8:e75461. doi: 10.1371/journal.pone.0075461

Louis, A., Crollius, H. R., and Robinson-Rechavi, M. (2012). How much does the amphioxus genome represent the ancestor of chordates? *Brief. Funct. Genomics* 11, 89–95. doi: 10.1093/bfgp/els003

Lu, T.-M., Luo, Y.-J., and Yu, J.-K. (2012). BMP and Delta/Notch signaling control the development of amphioxus epidermal sensory neurons: insights into the evolution of the peripheral sensory system. *Development* 139, 2020–2030. doi: 10.1242/dev.073833

Marlétaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564, 64–70. doi: 10.1038/s41586-018-0734-6

McDougall, A., Chenevert, J., Godard, B. G., and Dumollard, R. (2019). "Emergence of embryo shape during cleavage divisions," in *Evo-Devo: Non-Model Species in Cell and Developmental Biology Results and Problems in Cell Differentiation*, eds W. Tworzydlo and S. M. Bilinski (Cham: Springer International Publishing), 127–154. doi: 10.1007/978-3-030-23459-1_6

Morov, A. R., Ukizintambara, T., Sabirov, R. M., and Yasui, K. (2016). Acquisition of the dorsal structures in chordate amphioxus. *Open Biol.* 6:160062. doi: 10.1098/rsob.160062

Nishikawa, T. (2004). A new deep-water lancelet (Cephalochordata) from off Cape Nomamisaki, SW Japan, with a proposal of the revised system recovering the genus *Asymmetron*. *Zoolog. Sci.* 21, 1131–1136. doi: 10.2108/zsj.21.1131

Ogasawara, M. (2000). Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: insight into evolution of the thyroid gland. *Dev. Genes Evol.* 210, 231–242. doi: 10.1007/s004270050309

Olsson, R. (1983). Club-shaped gland and endostyle in larval *Branchiostoma lanceolatum* (Cephalochordata). *Zoomorphology* 103, 1–13. doi: 10.1007/BF00312054

Ono, H., Koop, D., and Holland, L. Z. (2018). Nodal and Hedgehog synergize in gill slit formation during development of the cephalochordate *Branchiostoma floridae*. *Development* 145:dev162586. doi: 10.1242/dev.162586

Pantzartzi, C. N., Pergner, J., Kozmikova, I., and Kozmik, Z. (2017). The opsin repertoire of the European lancelet: a window into light detection in a basal chordate. *Int. J. Dev. Biol.* 61, 763–772. doi: 10.1387/ijdb.170139zk

Paris, M., Escrivá, H., Schubert, M., Brunet, F., Brtko, J., Ciesielski, F., et al. (2008). Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr. Biol.* 18, 825–830. doi: 10.1016/j.cub.2008.04.078

Patalano, S., Prulière, G., Prodon, F., Paix, A., Dru, P., Sardet, C., et al. (2006). The aPKC–PAR-6–PAR-3 cell polarity complex localizes to the centrosome attracting body, a macroscopic cortical structure responsible for asymmetric divisions in the early ascidian embryo. *J. Cell Sci.* 119, 1592–1603. doi: 10.1242/jcs.02873

Poss, S. G., and Boschung, H. T. (1996). Lancelets (Cephalochordata: Branchiostomatidae): how many species are valid? *Isr. J. Zool.* 42, S13–S66. doi: 10.1080/00212210.1996.10688872

Prulière, G., Cosson, J., Chevalier, S., Sardet, C., and Chenevert, J. (2011). Atypical protein kinase C controls sea urchin ciliogenesis. *Mol. Biol. Cell* 22, 2042–2053. doi: 10.1091/mbc.e10-10-0844

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Richardson, M. K., and Wright, G. M. (2003). Developmental transformations in a normal series of embryos of the sea lamprey *Petromyzon marinus* (Linnaeus). *J. Morphol.* 257, 348–363. doi: 10.1002/jmor.10119

Sardet, C., McDougall, A., Yasuo, H., Chenevert, J., Pruliere, G., Dumollard, R., et al. (2011). "Embryological methods in ascidians: the Villefranche-sur-Mer protocols," in *Vertebrate Embryogenesis Methods in Molecular Biology*, ed. F. Pelegri (Totowa, NJ: Humana Press), 365–400. doi: 10.1007/978-1-61779-210-6_14

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Schubert, M., Meulemans, D., Bronner-Fraser, M., Holland, L. Z., and Holland, N. D. (2003). Differential mesodermal expression of two amphioxus *MyoD* family members (*AmphiMRF1* and *AmphiMRF2*). *Gene Expr. Patterns* 3, 199–202. doi: 10.1016/S1567-133X(02)00099-6

Shi, C. G., Huang, J., Chen, S. X., Li, G., and Wang, Y. Q. (2018). Generation of two transgenic amphioxus lines using the *Tol2* transposon system. *J. Genet. Genom.* 45, 513–516. doi: 10.1016/j.jgg.2018.06.002

Simakov, O., Marlétaz, F., Yue, J.-X., O'Connell, B., Jenkins, J., Brandt, A., et al. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nat. Ecol. Evol.* 4, 820–830. doi: 10.1038/s41559-020-1156-z

Slack, J. M. W., Holland, P. W. H., and Graham, C. F. (1993). The zootype and the phylotypic stage. *Nature* 361, 490–492. doi: 10.1038/361490a0

Smith, M. M., Smith, L. C., Cameron, R. A., and Urry, L. A. (2008). The larval stages of the sea urchin, *Strongylocentrotus purpuratus*. *J. Morphol.* 269, 713–733. doi: 10.1002/jmor.10618

Somorjai, I., Bertrand, S., Camasses, A., Haguenauer, A., and Escrivá, H. (2008). Evidence for stasis and not genetic piracy in developmental expression patterns of *Branchiostoma lanceolatum* and *Branchiostoma floridae*, two amphioxus species that have evolved independently over the course of 200 Myr. *Dev. Genes Evol.* 218, 703–713. doi: 10.1007/s00427-008-0256-6

Stokes, M. (1997). Larval locomotion of the lancelet. *J. Exp. Biol.* 200, 1661–1680. doi: 10.1242/jeb.200.11.1661

Stokes, M. D., and Holland, N. D. (1995). Embryos and larvae of a lancelet, *Branchiostoma floridae*, from hatching through metamorphosis: growth in the laboratory and external morphology. *Acta Zool.* 76, 105–120. doi: 10.1111/j.1463-6395.1995.tb00986.x

Su, L., Shi, C., Huang, X., Wang, Y., and Li, G. (2020). Application of CRISPR/Cas9 nuclease in amphioxus genome editing. *Genes* 11:1311. doi: 10.3390/genes11111311

Subirana, L., Farstey, V., Bertrand, S., and Escrivá, H. (2020). *Asymmetron lucayanum*: how many species are valid? *PLoS One* 15:e0229119. doi: 10.1371/journal.pone.0229119

Tung, T. C., Wu, S. C., and Tung, Y. Y. F. (1961). Experimental studies on neural induction in amphioxus. *Sci. Sin.* 7, 263–270.

Tung, T. C., Wu, S. C., and Tung, Y. Y. F. (1962). The presumptive areas of the egg of amphioxus. *Sci. Sin.* 11, 629–644.

Urata, M., Yamaguchi, N., Henmi, Y., and Yasui, K. (2007). Larval development of the Oriental lancelet, *Branchiostoma belcheri*, in laboratory mass culture. *Zool. Sci.* 24, 787–797. doi: 10.2108/zsj.24.787

Wickstead, J. H. (1967). *Branchiostoma lanceolatum* larvae: some experiments on the effect of thiouracil on metamorphosis. *J. Mar. Biol. Assoc.* 47, 49–59. doi: 10.1017/S0025315400033555

Willey, A. (1894). *Amphioxus and the Ancestry of the Vertebrates*. London: Macmillan and Co.

Wu, H.-R., Chen, Y.-T., Su, Y.-H., Luo, Y.-J., Holland, L. Z., and Yu, J.-K. (2011). Asymmetric localization of germline markers *Vasa* and *Nanos* during early development in the amphioxus *Branchiostoma floridae*. *Dev. Biol.* 353, 147–159. doi: 10.1016/j.ydbio.2011.02.014

Yang, K. Y., Chen, Y., Zhang, Z., Ng, P. K.-S., Zhou, W. J., Zhang, Y., et al. (2016). Transcriptome analysis of different developmental stages of amphioxus reveals dynamic changes of distinct classes of genes during development. *Sci. Rep.* 6:23195. doi: 10.1038/srep23195

Yong, L. W., Kozmikova, I., and Yu, J.-K. (2019). "Using amphioxus as a basal chordate model to study BMP signaling pathway," in *Bone Morphogenetic Proteins: Methods and Protocols Methods in Molecular Biology*, ed. M. B. Rogers (New York, NY: Springer), 91–114. doi: 10.1007/978-1-4939-8904-1_8

Yu, J.-K., and Holland, L. Z. (2009). Amphioxus whole-mount in situ hybridization. *Cold Spring Harb. Protoc.* 2009:db.rot5286. doi: 10.1101/pdb.prot5286

Yue, J.-X., Yu, J.-K., Putnam, N. H., and Holland, L. Z. (2014). The transcriptome of an amphioxus, *Asymmetron lucayanum*, from the Bahamas: a window into chordate evolution. *Genome Biol. Evol.* 6, 2681–2696. doi: 10.1093/gbe/evu212

Zhang, Q.-J. (2017). *Taxonomy of genus Branchiostoma in Xiamen waters and continuous breeding of two lancelets in the laboratory*. Doctoral dissertation. Xiamen: Xiamen University.

Zhang, Q.-J., Luo, Y.-J., Wu, H.-R., Chen, Y.-T., and Yu, J.-K. (2013). Expression of germline markers in three species of amphioxus supports a preformation mechanism of germ cell development in cephalochordates. *EvoDevo* 4:17. doi: 10.1186/2041-9139-4-17

Zhang, Q.-J., Sun, Y., Zhong, J., Li, G., Lü, X.-M., and Wang, Y.-Q. (2007). Continuous culture of two lancelets and production of the second filial generations in the laboratory. *J. Exp. Zool. B Mol. Dev. Evol.* 308B, 464–472. doi: 10.1002/jez.b.21172

Zhang, Q.-J., Zhong, J., Fang, S.-H., and Wang, Y.-Q. (2006). *Branchiostoma japonicum* and *B. belcheri* are distinct lancelets (Cephalochordata) in Xiamen waters in China. *Zool. Sci.* 23, 573–579. doi: 10.2108/zsj.23.573

Zhu, X., Shi, C., Zhong, Y., Liu, X., Yan, Q., Wu, X., et al. (2020). Cilia-driven asymmetric Hedgehog signalling determines the amphioxus left-right axis by controlling *Dand5* expression. *Development* 147:dev182469. doi: 10.1242/dev.182469

Zieger, E., Garbarino, G., Robert, N. S. M., Yu, J.-K., Croce, J. C., Candiani, S., et al. (2018). Retinoic acid signaling and neurogenic niche regulation in the developing peripheral nervous system of the cephalochordate amphioxus. *Cell. Mol. Life Sci.* 75, 2407–2429. doi: 10.1007/s00018-017-2734-3

Zieger, E., Lacalli, T. C., Pestarino, M., Schubert, M., and Candiani, S. (2017). The origin of dopaminergic systems in chordate brains: insights from amphioxus. *Int. J. Dev. Biol.* 61, 749–761. doi: 10.1387/ijdb.170153sc

Conflict of Interest: 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 © 2021 Carvalho, Lahaye, Yong, Croce, Escrivá, Yu and Schubert. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 May 2021
doi: 10.3389/fcell.2021.639522

Prenatal Developmental Trajectories of Fluctuating Asymmetry in Bat Humeri

Camilo López-Aguirre[1,2]*, *Suzanne J. Hand*[2], *Daisuke Koyabu*[3,4], *Vuong Tan Tu*[5,6] *and Laura A. B. Wilson*[2,7]

[1] *Department of Anthropology, University of Toronto Scarborough, Toronto, ON, Canada,* [2] *Earth and Sustainability Science Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia,* [3] *Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Kowloon, Hong Kong,* [4] *Research and Development Center for Precision Medicine, University of Tsukuba, Tsukuba, Japan,* [5] *Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, Hanoi, Vietnam,* [6] *Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam,* [7] *School of Archaeology and Anthropology, The Australian National University, Canberra, ACT, Australia*

Fluctuating asymmetry (random fluctuations between the left and right sides of the body) has been interpreted as an index to quantify both the developmental instabilities and homeostatic capabilities of organisms, linking the phenotypic and genotypic aspects of morphogenesis. However, studying the ontogenesis of fluctuating asymmetry has been limited to mostly model organisms in postnatal stages, missing prenatal trajectories of asymmetry that could better elucidate decoupled developmental pathways controlling symmetric bone elongation and thickening. In this study, we quantified the presence and magnitude of asymmetry during the prenatal development of bats, focusing on the humerus, a highly specialized bone adapted in bats to perform under multiple functional demands. We deconstructed levels of asymmetry by measuring the longitudinal and cross-sectional asymmetry of the humerus using a combination of linear measurements and geometric morphometrics. We tested the presence of different types of asymmetry and calculated the magnitude of size-controlled fluctuating asymmetry to assess developmental instability. Statistical support for the presence of fluctuating asymmetry was found for both longitudinal and cross-sectional asymmetry, explaining on average 16% of asymmetric variation. Significant directional asymmetry accounted for less than 6.6% of asymmetric variation. Both measures of fluctuating asymmetry remained relatively stable throughout ontogeny, but cross-sectional asymmetry was significantly different across developmental stages. Finally, we did not find a correspondence between developmental patterns of longitudinal and cross-sectional asymmetry, indicating that processes promoting symmetrical bone elongation and thickening work independently. We suggest various functional pressures linked to newborn bats' ecology associated with longitudinal (altricial flight capabilities) and cross-sectional (precocial clinging ability) developmental asymmetry differentially. We hypothesize that stable magnitudes of fluctuating asymmetry across development could indicate the presence of developmental mechanisms buffering developmental instability.

Keywords: fluctuating asymmetry, Chiroptera, prenatal development, humerus, morphogenesis

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Shin-ichi Fujiwara,
Nagoya University, Japan
Benedikt Hallgrimsson,
University of Calgary, Canada

Correspondence:
Camilo López-Aguirre
c.lopezaguirre@utoronto.ca

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *09 December 2020*
Accepted: *14 April 2021*
Published: *26 May 2021*

Citation:
López-Aguirre C, Hand SJ, Koyabu D, Tu VT and Wilson LAB (2021) Prenatal Developmental Trajectories of Fluctuating Asymmetry in Bat Humeri. Front. Cell Dev. Biol. 9:639522. doi: 10.3389/fcell.2021.639522

INTRODUCTION

Bats (Chiroptera) have a highly specialized postcranial morphology adapted to perform self-powered flight, a trait unique among mammals. Compared to terrestrial mammals, the forelimb skeleton of bats is characterized by aerial-based functional adaptations that provide structural support to the wing (bone elongation of distal bones) (Panyutina et al., 2015), withstand biomechanical loading (increased bone density) (Swartz, 1997; Dumont, 2010), enhance flight maneuverability (proximo-distal demineralization) (Papadimitriou et al., 1996), and facilitate the generation of flight power (increase of muscle mass in the shoulder girdle) (Panyutina et al., 2015). Moreover, morphological adaptations related to different functional pressures can occur within a single bone (e.g., increased muscle attachment area and biomechanical load resistance in the humerus; Panyutina et al., 2015). Such a variety of long bone adaptations arise ontogenically, mainly through two bone remodeling processes: bone elongation and bone thickening (Enlow, 1963). The former is linked to axial bone deposition *via* endochondral ossification at growth plates (Webber et al., 2015), whereas the latter results from endosteal and periosteal bone deposition *via* intramembranous ossification in the periosteum/endosteum (Kronenberg, 2003). These two processes derive from independent morphogenetic pathways that respond to different factors, indicating decoupled developmental trajectories controlling the optimal development of individual forelimb bones (Atchley and Hall, 1991; Montoya-Sanhueza et al., 2019). However, our understanding of the differences in developmental trajectories of bone thickening and elongation in the bat wing remains largely limited.

Uncovering the historical trajectories that led to the morphological diversification and specialization of bats has been greatly limited by a markedly incomplete fossil record (Brown et al., 2019). Evolutionary developmental biology has emerged as a promising approach to study the evolution of bats while circumventing limitations in the fossil record (Adams, 2008; Cooper et al., 2012; Camacho et al., 2019; López-Aguirre et al., 2019,a,b). Studies have provided evidence for the ontogenetic mechanisms behind forelimb specialization in bats and the evolution of vertebrate flight (Sears et al., 2006; Adams, 2008; Cretekos et al., 2008; Farnum et al., 2008; Cooper et al., 2012; Adams and Shaw, 2013), ecology-driven deviations in chiropteran development from general mammalian patterns, and the phylogenetic signal in postcranial development (Adams, 1992; Koyabu and Son, 2014; López-Aguirre et al., 2019,a,b).

For body plans that are naturally symmetrical, deviations from an "ideal" state have been interpreted as a signal of reduced fitness (Dongen, 2006). Accordingly, quantifying the deviation from perfect symmetry can indicate the amount of stress an organism undergoes and its homeostatic capacity (i.e., buffering of instabilities to maintain fitness) (Gummer and Brigham, 1995; Aparicio and Bonal, 2002). The main regulatory mechanisms that influence phenotypic symmetry occur ontogenically, when the genotypic and phenotypic mechanisms involved in morphogenesis can be destabilized by genetic or environmental stressors (i.e., developmental noise) (Hallgrímsson, 1998, 1999; Hallgrímsson et al., 2003; Kellner and Alford, 2003). Evolutionary studies have provided evidence for the heritability of an organism's capacity to buffer developmental noise (developmental stability/instability, DI), suggesting that natural selection can act as a regulator of phenotypic asymmetry (Tocts et al., 2016).

Phenotypic asymmetry in animals with bilateral asymmetry can be quantified by computing the morphometric differences between the right and the left side of the body (right–left, R–L) (Palmer, 1994). Advances in theoretical framework to quantify body asymmetry has enabled the recognition of three types of asymmetry, each with a different biological interpretation: fluctuating asymmetry (FA) is characterized as random deviations from "ideal" perfectly symmetrical phenotypes, directional asymmetry (DA) is described as a natural tendency to have consistently asymmetrical phenotypes (i.e., one side always larger than the other), and antisymmetry (AS) represents a pattern where symmetrical phenotypes are least favored and asymmetry is equally distributed across both sides (Klingenberg, 2015). All three types of asymmetry have also been described in mathematical terms: FA is characterized by a normal distribution of asymmetry values (R–L) along a value mean of zero, DA is described by a normal distribution of asymmetry values along a mean different to zero, and AS is identified where asymmetry values have a bimodal distribution and most values are different from zero (Palmer, 1994; Klingenberg, 2015). Combining the biological and mathematical interpretations of phenotypic asymmetry, FA has been regularly used as a possible indicator of DI, although some studies argue that DA and AS can also be indicators of DI (Palmer, 1994; Leamy and Klingenberg, 2005). Despite the utility of studying FA and DI, the efficacy of the theoretical framework traditionally applied to detect a real FA–DI link has been debated (Palmer, 1994; Dongen, 2006; Klingenberg, 2015).

The FA–DI link can vary in response to ecological, genetic, environmental, and developmental factors, stressing the need to study it at multiple scales to test a variety of hypotheses (Kellner and Alford, 2003). Swaddle and Witter (1997) summarized and Kellner and Alford (2003) postulated and tested predictions for a list of hypotheses on the ontogeny of FA describing developmental patterns of asymmetry and possible evolutionary mechanisms shaping them. These include the following: small fluctuations during early growth are magnified during later morphogenesis (i.e., magnification of asymmetry hypothesis), side-biased environmental influences can skew growth toward asymmetrical phenotypes (i.e., directional external cues hypothesis), accumulative growth of independent subunits will homogenize morphogenesis, reducing asymmetry throughout development (i.e., coin toss hypothesis), and developmental feedback mechanisms will stabilize asymmetric growth between structures by either promoting or constraining growth (i.e., compensatory growth hypothesis). Studies on the developmental basis of FA have also been restricted in scope (mostly focused on postnatal development) and study groups (invertebrates and captive populations) (Hallgrímsson, 1999; Hallgrímsson et al., 2003; Leamy and Klingenberg, 2005; Blackburn, 2011; Perchalski et al., 2018), limiting our understanding of variation in wild non-model species and potential insights into the mechanisms controlling FA in early development. Based on the limited studies available, decreasing magnitudes of FA across prenatal development have been reported as a result of the interaction between variations in timing of growth and growth rates (Hallgrímsson, 1998; Hallgrímsson et al., 2003; Kellner and Alford, 2003). In contrast, magnitudes of FA have been reported to increase during the development of the mammalian skeleton due to cumulative variability in growth regulation and/or bone remodeling (i.e., morphogenetic drift model; Hallgrímsson, 1998).

Asymmetry in bats has been studied at the cranial and postcranial level (Juste et al., 2001a,b; Lüpold et al., 2004; Voigt et al., 2005; López-Aguirre and Pérez-Torres, 2015), all studies being based on the analysis of adult specimens only. Forelimb FA has been associated with differential reproductive success (Lüpold et al., 2004; Voigt et al., 2005), suggesting that sexual selection favors symmetric individuals in *Saccopteryx bilineata* (Voigt et al., 2005) and a significant link between asymmetry and reproductive potential in *Carollia perspicillata* (Monteiro et al., 2019). A correlation between wing FA and resistance to environmental stress and resilience to anthropogenic habitat change has also been assessed, suggesting high resilience in Neotropical bat species (de Figueiredo et al., 2015; Castillo-Figueroa, 2018). Compensatory growth has been reported in the wing of the vampire bat *Desmodus rotundus* as a way to maintain wingspan symmetry (Ueti et al., 2015), while sex-based differences in the magnitudes of FA have been reported in the wing of *D. rotundus* and the cranium of *Artibeus lituratus* (López-Aguirre and Pérez-Torres, 2015; Ueti et al., 2015). Research also indicates that variation in the levels of FA across morphological (both cranial and postcranial) traits could depend on functional importance, with FA decreasing in traits under higher functional demands for feeding and locomotion (Gummer and Brigham, 1995; López-Aguirre and Pérez-Torres, 2015; Robaina et al., 2017). Despite the repeated study of bat forelimb FA, its ontogenetic basis and how it varies across different bone growth dimensions (i.e., bone elongation and thickening) remain unknown.

The objective of this study was to assess whether the ontogenetic trajectories of phenotypic asymmetry during bone elongation and thickening in bats are decoupled. We analyzed the presence and magnitude of FA in the prenatal morphogenesis of the humerus in bats, representing the first developmental study of FA in Chiroptera. We focused on the humerus because it represents a clear example of multiple functional demands acting on a single bone (i.e., withstanding torsional and bending stress, increasing muscle insertion area, and controlling the maneuverability of the wing) (Swartz et al., 1992; Panyutina et al., 2015). Humeral cross-sectional shape has been found to reflect foraging differences across bat taxa (López-Aguirre et al., 2019). Furthermore, prenatal limb FA has been described as an accurate indicator of DI in human fetuses (Klingenberg and Nijhout, 1999; Broek et al., 2017). We quantify asymmetry based on bone elongation and cross-sectional cortical bone deposition as a way to exemplify the multipatterned process of bone growth (Klingenberg and McIntyre, 1998). Humeral length asymmetry is commonly used in bat studies (Gummer and Brigham, 1995; Voigt et al., 2005; de Figueiredo et al., 2015; Ueti et al., 2015; Robaina et al., 2017; Castillo-Figueroa, 2018), whereas cross-sectional asymmetry is commonly measured in other mammals (Macintosh et al., 2013; Wilson and Humphrey, 2015; Perchalski et al., 2018). Based on previous studies inferring compensatory growth in the wing of bats and previous evidence of decreasing FA across development in terrestrial mammals (Hallgrímsson et al., 2003; Ueti et al., 2015), we hypothesize that asymmetry of the humerus will decrease throughout ontogeny. Because bone thickening and elongation in mammals represent independent morphogenetic trajectories, we predict that cross-sectional and longitudinal asymmetry will not be correlated.

MATERIALS AND METHODS

Sampling

A total of 66 prenatal specimens from 11 bat species (**Table 1**, see **Figures 1A–C**) were collected through taxonomic fieldwork in Vietnam by VTT and DK under collection permit no. 972/UBND-TH issued by Tuyen Quang Provincial People's Committee and research and ethics permit no. 322/STTNSV of the Institute of Ecology and Biological Resources, Vietnam Academy of Sciences. All specimens were fixed in Serra's fixative (ethanol, formalin, and glacial acetic acid mixed 6:3:1 by volume) for 48 h, then transferred, and preserved in 70% ethanol. 3D scanning of the embryos and fetuses was performed using a microfocal X−ray computed tomography system at the University Museum, University of Tokyo (TXS225−ACTIS; TESCO; Tokyo, Japan), with 70-kV source voltage and 114 μA source currents at a resolution of 36 μm. All osseous skeletal elements were segmented by the first author using the thresholding tool and the predetermined bone setting in MIMICS v. 20 software (Materialise NV, Leuven, Belgium). To standardize finer manual segmentation in early embryos, thresholding of Hounsfield unit values of osseous tissue was performed using the half-maximum

TABLE 1 | Composition of the sample used in this study, including the number of specimens and the developmental stages per species.

Family	Species	N	Developmental stages
Hipposideridae	*Aselliscus dongbacana*	14	4, 6–10
Hipposideridae	*Aselliscus stoliczkanus*	13	4–5, 7–10
Pteropodidae	*Cynopterus sphinx*	4	2, 8, 10
Vespertilionidae	*Hesperoptenus blandfordi*	5	1, 4–5, 8
Hipposideridae	*Hipposideros larvatus*	4	7–9
Vespertilionidae	*Kerivoula hardwickii*	7	2–5
Miniopteridae	*Miniopterus schreibersii*	2	10
Vespertilionidae	*Myotis* sp.	2	4–5
Rhinolophidae	*Rhinolophus pearsonii*	1	4
Rhinolophidae	*Rhinolophus pusillus*	1	2
Rhinolophidae	*Rhinolophus thomasi*	10	1, 6–9

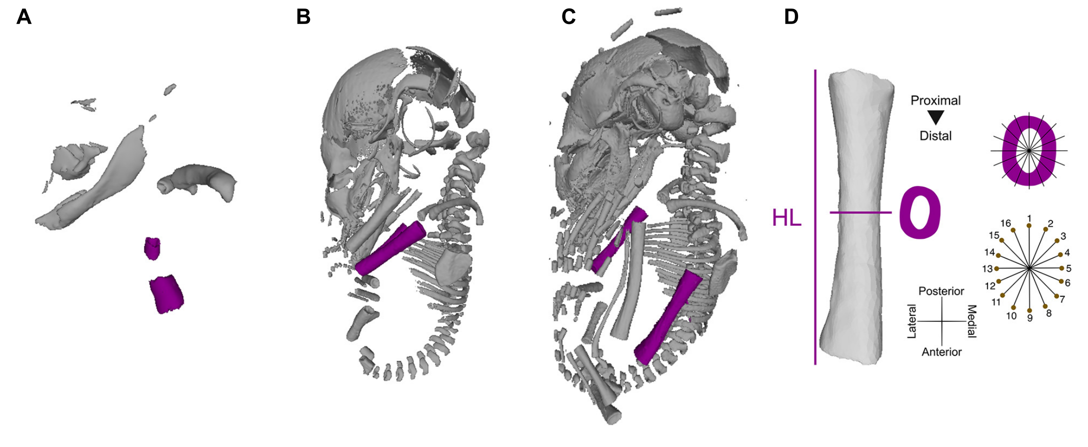


FIGURE 1 | 3D virtual models of ontogenetic series of *Rhinolophus thomasii*, representing the samples from which raw measurements were taken. Stage 1 **(A)**, stage 6 **(B)**, and stage 9 **(C)**. Landmarking and measuring protocol used to quantify humeral length (HL) and cortical thickness **(D)**.

height method (i.e., gradual change in computerized tomography values at the boundary of a structure) (Spoor et al., 1993). This method homogenously retrieved diaphyseal osseous tissue only, removing segmenting inconsistencies related to the imaging of diaphyseal cartilaginous tissue. Smoothing techniques were not used to prevent artificially changing the dimensions of the periosteal surface of the models. After selecting specimens with at least partially ossified and unbroken humeri, three individuals were excluded from the final sample. The left and right humeri were segmented and exported as STL files for further processing and analysis.

Data Collection

Humeral development was described based on bone elongation (i.e., humeral length, HL) and cortical bone deposition (i.e., periosteal diameter as a proxy for cross-sectional cortical thickness, CT) (**Figure 1D**). To estimate cross-sectional and HL measurements, 3D humeri models were imported into Rhinoceros 5.0 (Robert McNeel & Associates, Seattle, WA). To remove the non-shape effects of translation, rotation, and scale, all humeri models were aligned to a standard position in 3D space, following a protocol for long-bone cross-sectional geometry (see Wilson and Humphrey, 2015). HL was automatically measured as the length of the long axis of a rectilinear box (i.e., bounding box) enclosing the model created using the *BoundingBox* command in Rhinoceros 5.0, preventing human measurement error (ME). Cross-sections of the left and right humeri were extracted at the midshaft (i.e., 50% of HL) for a total of 126 cross-sections [63 × 2 (left and right sides)]. The midshaft was extracted by placing a perpendicular axis intersecting the 3D model at the midpoint of the length of the rectilinear bounding box, dividing HL in half. Small deviations in the placing of the cross-sections could occur due to asymmetrical development of cartilaginous epiphyseal tissue. Humeral CT was quantified using a geometric morphometrics-based approach. Following the method described in Wilson and Humphrey (2015), we used a set of 16 equiangular landmarks semi-automatically placed along the periosteal surface of each cross-section (**Figure 1D**). The cross-sections were aligned with a set of 16 equiangular radii along the centroid of each cross-section, aligning the radii with the anatomical axes of the bone (e.g., radii 1 and 9 represent the antero-posterior axis). The landmarks were automatically placed on the intersection of individual radii and the periosteal surface of each cross-section. CT was quantified as the average of interlandmark distances between pairs of landmarks that formed linear axes (e.g., landmarks 1 and 9; see **Figure 1D**) using the interlmkdist function in R package Geomorph 3.2 (Adams et al., 2013). Our automated cross-sectional geometric morphometrics-based approach enables the quantification of CT while circumventing the lack of identifiable homologous landmarks early in prenatal development. To control for the effect of matching bilateral symmetry (i.e., the left and right sides of the body are mirror images), the landmark coordinates of the cross-sections of the right side were reflected along the antero-posterior axis by multiplying the coordinates of that axis by −1. Given the lack of Carnegie staging systems for many non-model taxa (nine of the 11 species in our sample), staging of developmental series was based on crown-to-rump length and bone ossification sequence as described in López-Aguirre et al. (2019a) and following the general patterns described in bat development (Cretekos et al., 2005). All developmental stages were represented by at least three individuals (**Supplementary Table 1**).

Estimation of Asymmetry

Asymmetry in bilateral organisms can be described as the difference between both sides of the body (e.g., right–left) (Palmer, 1994), which we estimated based on HL (longitudinal asymmetry) and CT (cross-sectional asymmetry). Individual longitudinal and cross-sectional asymmetries were quantified as the signed difference between right and left HL, negative and positive values indicating the directionality of asymmetry (Palmer, 1994). Individual longitudinal FA was described as the size-corrected signed difference between sides, and per stage FA was calculated as the variance of individual FA measurements across a population (i.e., FA6 index) (Palmer, 1994). FA6 is a single-trait index that expresses FA as the variance in asymmetry proportional to trait size (i.e., HL) in an individual so as to truly represent DI and not developmental bone growth,

$$\mathrm{var}\left[\frac{R-L}{0.5*(RL)}\right].$$

Data Analysis

The presence of FA and DA in longitudinal and cross-sectional asymmetry was statistically tested with full-factorial ANOVAs using side (left or right), individual, and duplicate as factors (FA $\sim$ side + individual + side/individual; see **Table 2**; Monteiro et al., 2019; Rivera and Neely, 2020). The side factor provides a statistical test for DA, whereas the side–individual interaction provides statistical tests for FA. Measurement error was not computed because the automated protocol implemented to obtain HL and CT ensures that no human error could affect the measuring process.

Developmental trajectories of individual longitudinal and cross-sectional FA were explored using box plots, while per stage FA6 was explored using bar plots. Statistical differences across developmental stages were tested using two ANOVAs, using the stage of each individual as a factor and unsigned longitudinal and cross-sectional FA (longitudinal FA$\sim$stage and cross-sectional FA$\sim$stage).

We tested the association of longitudinal and cross-sectional FA across development using a linear regression model based on ordinary least squares. We used linear regression models to test the effect of the number of specimens and species in values of average longitudinal and cross-sectional FA per developmental stage to assess whether uneven sample composition affected our results. Using a subsample of the best sampled genus, we explored the association between peaks in humeral growth and peaks in magnitudes of asymmetry. Finally, we tested the consistency of our results by recreating the ontogenetic patterns of longitudinal and cross-sectional FA in a subsample of the best sampled species (*Aselliscus dongbacana* and *Aselliscus stoliczkanus*).

TABLE 2 | ANOVA statistical tests of significance of fluctuating asymmetry (FA) and directional asymmetry (DA) in cortical thickness (CT) and humeral length (HL).

	Df	SS	MS	Rsq	*F*	Pr (> *F*)
HL						
Individual	62	1063.9	17.160	0.9925	2304	<0.0001
Side	1	0	0	0	0	0.983
Individual*Side	62	0.8	0.013	0.0075	1.759	<0.05
Residuals	125	0.9	0.007	4.8E-30		
CT						
Individual	62	2.472E-05	3.987E-07	0.5511	6.484E+23	<0.0001
Side	1	9.500E-08	9.550E-08	0.1320	1.553E+23	<0.0001
Individual*Side	62	1.422E-05	2.293E-07	0.3169	3.730E+23	<0.0001
Residuals	125	0	0	0		

*Side factor tests for DA, individual*side interaction tests for FA, and replicate tests for ME. Df, degrees of freedom; SS, sum of squares; MS, mean square; Rsq, R square.*

RESULTS

Individual Longitudinal and Cross-Sectional Asymmetry

Distribution of individuals' values of longitudinal and cross-sectional FA demonstrate a normal distribution with a mean near zero (**Figure 2**), supporting the presence of FA in longitudinal and cross-sectional humeral asymmetry throughout development. Moreover, 42.86% of individuals had negative values of longitudinal FA (**Figure 2A**), indicating that a narrow majority of individuals had larger right humeri (57.14%). Furthermore, 39.68% of individuals showed negative values of cross-sectional asymmetry (**Figure 2B**). Across datasets (cross-sectional and HL), only three individuals were found to have perfectly symmetrical humeri (i.e., R–L = 0) for HL in developmental stages 5, 7, and 8.

Box plots of individuals' longitudinal and cross-sectional FA values illustrated independent trajectories of cross-sectional and longitudinal humeral FA across development (**Figure 3**). Unsigned longitudinal FA indicated two peaks of high values of longitudinal FA in individuals early in prenatal development (stages 2, 3, and 6), separated by a sharp decrease in longitudinal FA values between stages 4 and 5. Longitudinal FA values steadily decreased from stage 7 onwards, with a noticeable increase in dispersion in stages 2 and 6 (**Figure 3A**). A similar pattern was found in our subsample of the best sampled species, with relatively stable longitudinal FA after stage 1 and an increase in variability in stages 8 and 9 (**Supplementary Figure 1**). Cross-sectional FA showed low magnitudes of FA in individuals in the early stages of prenatal development (stages 1 and 2), followed by an increase between stages 3 and 5 (**Figure 3B**). Individuals in stage 5 showed the highest dispersion of FA values. Cross-sectional FA consistently decreased from stage 6 onwards. The cross-sectional FA in our subsample remained relatively stable after stage 3, followed by a decrease from stage 7, similar to our general results (**Supplementary Figure 1**).

Presence and Magnitude of Fluctuating Asymmetry

The ANOVA of asymmetry on HL supported the presence of both DA and FA in longitudinal asymmetry (**Table 2**). Nevertheless, FA accounted for only 0.007% of variation, DA for less than 0.0001% of the variation, and individual variation for 99.25% of variability in raw HL. The markedly low values of longitudinal FA suggest that the extraction of the cross-sections did not vary due to asymmetrical growth rates between epiphyseal plates. The ANOVA of cross-sectional asymmetry found statistical support for the presence of FA and DA (**Table 2**). DA accounted for only 13.20% of cross-sectional variation and FA for 31.69% of variation, respectively. Individual variation explained the highest

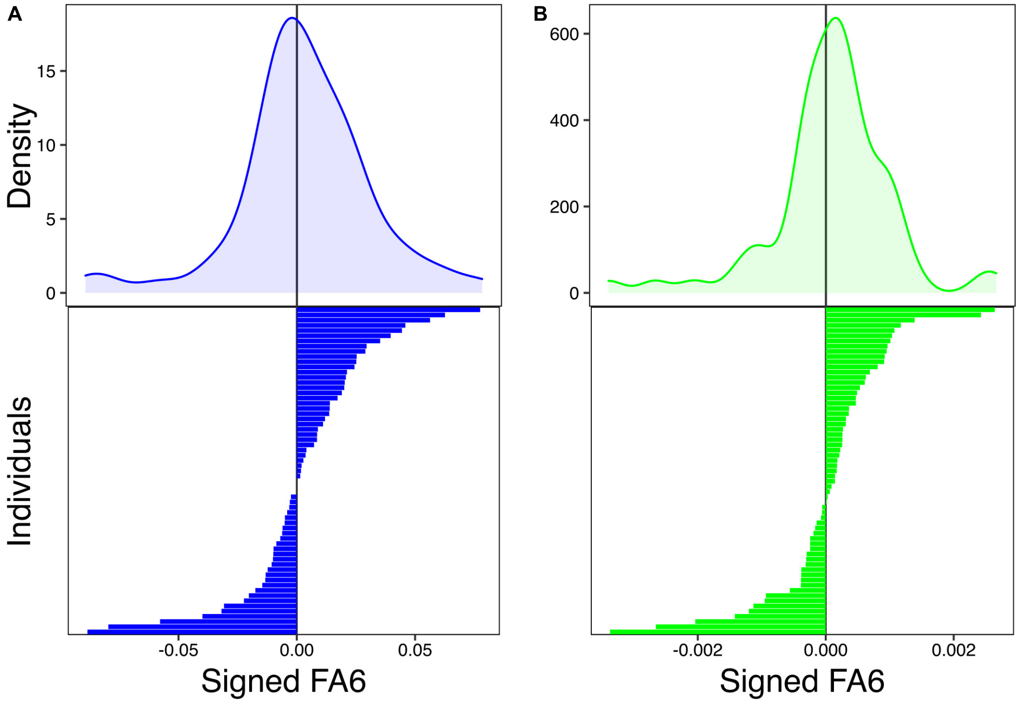


FIGURE 2 | Probability density functions of distribution of values of longitudinal (signed fluctuating asymmetry, FA) **(A)** and cross-sectional FA (signed FA) **(B)**. Bar plots show the individual asymmetry values from which density functions were estimated.

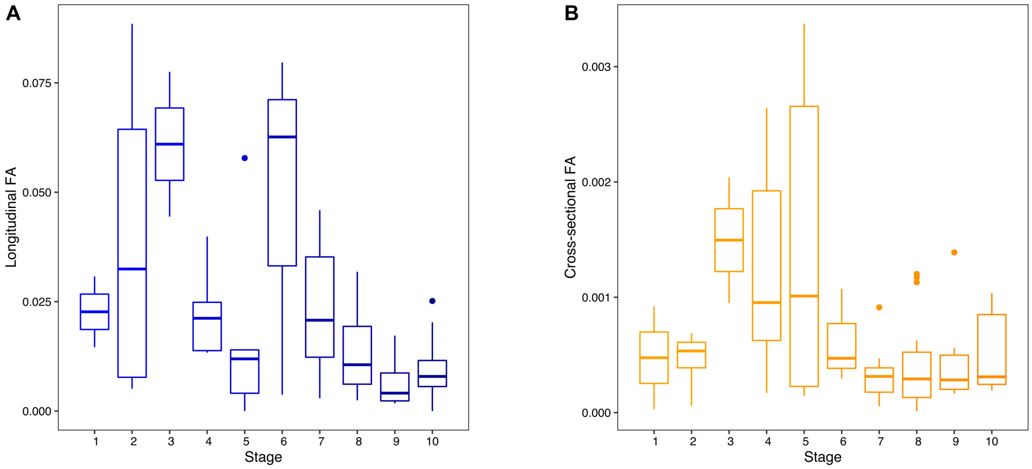


FIGURE 3 | Box plots of individual longitudinal (unsigned fluctuating asymmetry, FA) **(A)** and cross-sectional (unsigned FA) **(B)** FA values across development. Stages 1–10 represent early to late prenatal development.

proportion of variation, accounting for 55.11% of cross-sectional asymmetrical variation.

Developmental Trajectories of Fluctuating Asymmetry

The per-stage longitudinal and cross-sectional FA (FA6) indicated disparate trajectories across development (**Figure 4**). The longitudinal FA values showed two peaks (**Figure 4A**): one early in development (stage 2) and another in intermediate prenatal development (stage 6). The lowest longitudinal FA values were found in stages 4 and 8–10. Cross-sectional FA showed a single peak during mid-prenatal development (stage 5), followed by a significant decrease from stages 6 to 10 (**Figure 4B**). Despite the differences in the trajectories of longitudinal and cross-sectional FA throughout development, our results indicate a trend of stable FA throughout development. The ANOVAs for differences in longitudinal and cross-sectional FA across

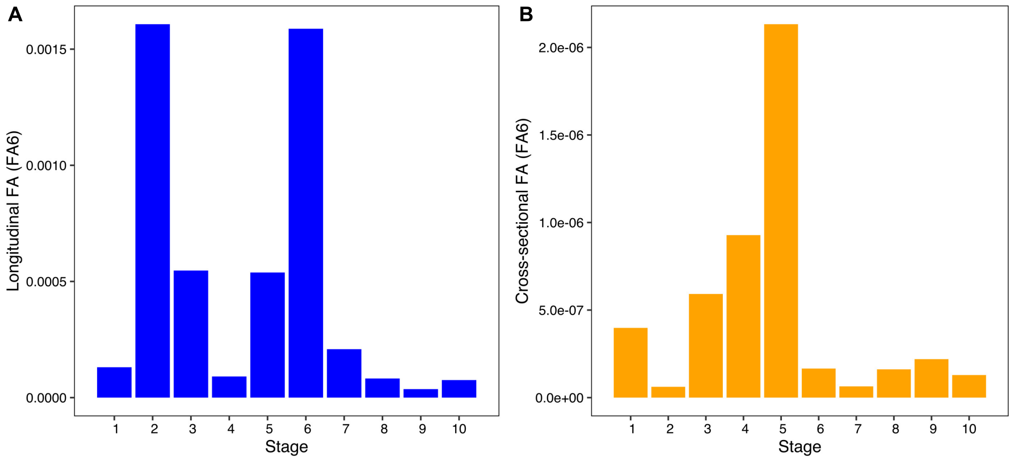


FIGURE 4 | Bar plot of per stage longitudinal **(A)** and cross-sectional **(B)** FA6 values across development. Stages 1–10 represent early to late prenatal development.

TABLE 3 | ANOVA test results for statistically significant differences in longitudinal and cross-sectional humeral fluctuating asymmetry (FA) across developmental stages.

	Df	SS	MS	Rsq	*F*	Pr (>*F*)
Longitudinal FA	1	0	0	−0.0164	0	0.993
Cross-sectional FA	1	2.1303E-06	2.1303E-06	0.0701	4.598	0.035

Df, Degrees of freedom; SS, Sum of squares; MS, Mean square; Rsq, R square.

developmental stages found statistically significant differences in cross-sectional FA but not in longitudinal FA (**Table 3**). Our results do not reveal clear similarities between peaks of FA and peaks of humeral growth, indicating that asymmetry does not increase when growth rates are higher (**Supplementary Figure 2**). Nevertheless, future studies should further explore the possible link between growth rates and magnitudes of asymmetry during development.

The scatterplots of longitudinal and cross-sectional asymmetry did not indicate a clear association between cross-sectional and longitudinal FA throughout development, with a slightly negative tendency (**Figure 5**). A linear regression model of cross-sectional and longitudinal FA confirmed a non-significant negative correlation between both dimensions of humeral FA (R^2 = 0.016, P = 0.979). The cross-sectional and longitudinal FA also did not correlate with the number of specimens and species per stage (cross-sectional FA: specimens R^2 = 0.136, P = 0.159 and species R^2 = 0.183, P = 0.121; longitudinal FA: specimens R^2 = 0.096, P = 0.66 and species R^2 = 0.072, P = 0.546), indicating that sampling heterogeneity did not influence our results.

DISCUSSION

This study is the first to demonstrate the presence of FA and DA during bat prenatal development. Our results showed stable magnitudes of longitudinal and cross-sectional FA across prenatal development, rejecting our hypothesis of decreasing asymmetry throughout ontogeny. However, our results also show that both measures of FA (longitudinal and cross-sectional) did not correlate across development, showing decoupled ontogenetic trajectories, similar to our prediction. Decoupled longitudinal and cross-sectional asymmetry magnitudes also indicate that the developmental pathways regulating symmetrical bone elongation and thickening likely act independently within a single bone. Signaling pathways regulating symmetric growth have been detected in the apical ectodermal ridge and the zone of polarizing activity of the limb bud (Allard and Tabin, 2009; Wolpert, 2010). Sonic hedgehog and fibroblast growth signaling factors play a major role during limb growth, controlling limb polarity (i.e., anterior–posterior limb differentiation), cell proliferation, and symmetric growth (Allard and Tabin, 2009; Wolpert, 2010). Previous studies have found decreasing trajectories of asymmetry across prenatal development in the postcranium of humans and mice, suggesting an interplay between the timing of morphogenesis and growth rates over the course of ontogeny (Hallgrímsson, 1998; Hallgrímsson et al., 2003). The compensatory growth between left and right sides of the body in response to increased asymmetry has been discussed as a mechanism to reduce DI (Aparicio, 1998; Ueti et al., 2015).

FA in bats has been an area of increasing interest in recent decades (Gummer and Brigham, 1995; Juste et al., 2001a,b; Lüpold et al., 2004; Voigt et al., 2005; de Figueiredo et al., 2015; López-Aguirre and Pérez-Torres, 2015; Ueti et al., 2015; Robaina et al., 2017; Castillo-Figueroa, 2018; Monteiro et al., 2019). A set of two studies explored the patterns of FA in insular populations of fruit bat species *Eidolon helvum* and *Rousettus egyptiacus* of the Gulf of Guinea (Juste et al., 2001a,b). Juste et al. (2001a) found similar patterns of FA across populations of both species, discussing the interpretation of population asymmetry parameters (i.e., consistent patterns of FA for a set of characters across populations of the same species), and suggested its scalability above the species level. Juste et al. (2001b) found consistent magnitudes of multivariate FA and a significant integration of asymmetry across traits and species, hypothesizing

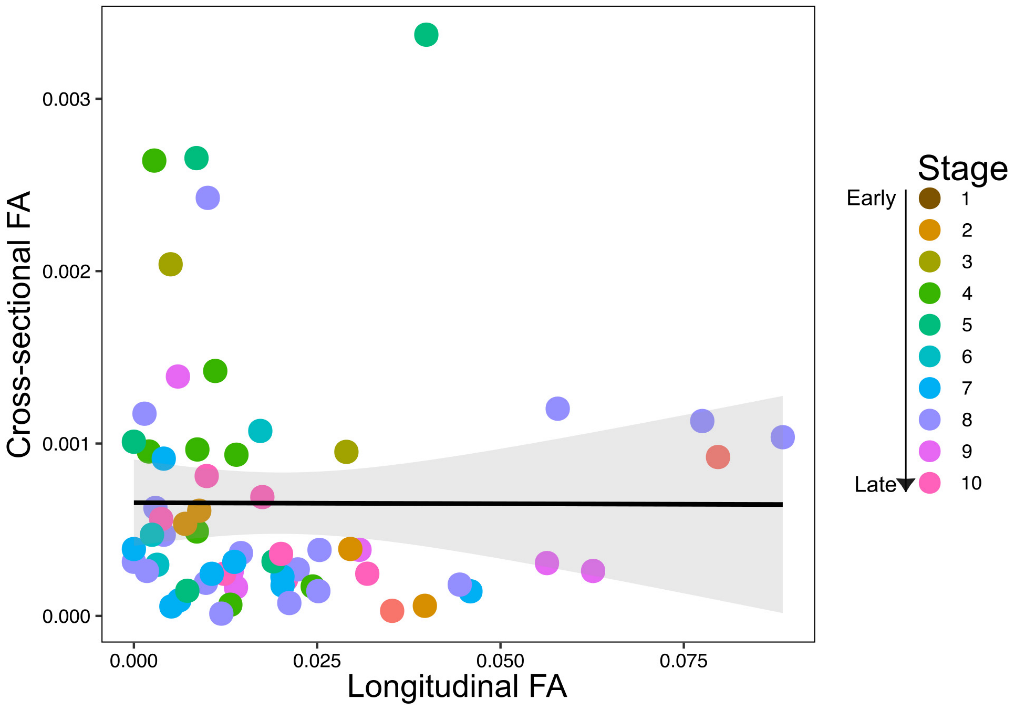


FIGURE 5 | Scatterplot of association between longitudinal (unsigned fluctuating asymmetry, FA) and cross-sectional (unsigned FA) fluctuating asymmetry across prenatal development. Colored dots represent developmental stages, with stages 1–10 representing early to late prenatal development.

high canalization in the developmental pathways controlling phenotypic asymmetry that are shared between the two species.

The presence and magnitude of FA has also been associated with reproductive success and sexual selection in bats (Voigt et al., 2005; Monteiro et al., 2019). Analyzing forearm length asymmetry, it has been suggested that sexual selection favors more symmetrical males in the polygynous greater sac-winged bat *Saccopteryx bilineata*, canalizing FA (Voigt et al., 2005). The number of offspring produced by males of *S. bilineata* was found to significantly decrease with increasing forearm asymmetry (Voigt et al., 2005). Increases in forearm asymmetry in the Neotropical frugivorous bat *Carollia perspicillata* have also been linked with a significant decrease in survival probability and the probability of more than one pregnancy per reproductive season (Monteiro et al., 2019). However, Lüpold et al. (2004) did not find a significant association between FA and other measures of individual fitness and allometry in the bat genitalia. All previous studies analyzing humeral asymmetry in bats were based on adult specimens with fully ossified epiphyses, whereas our study could not analyze cartilaginous osseous tissue.

Studies have also linked the presence and magnitude of FA to anthropogenic perturbations and habitat degradation, analyzing four Neotropical phyllostomid species (i.e., *Artibeus lituratus*, *Artibeus planirostris*, *C. perspicillata*, and *Sturnira lilium*) (de Figueiredo et al., 2015; Castillo-Figueroa, 2018). Neither study explicitly tested the presence of the three kinds of asymmetry, casting doubt on their interpretation that FA is an accurate index of resilience and adaptation of species to perturbations. Robaina et al. (2017) may have also insufficiently assessed the presence and magnitude of all types of asymmetry before drawing conclusions on the validity of FA to reflect functional importance. Our results suggest that special attention to the statistical framework used to describe the biological and theoretical interpretation of asymmetry in bats is warranted.

Asymmetry across multiple traits has shown decoupled patterns, reflecting functional differences and indicating independent developmental mechanisms controlling phenotypic symmetry across different structures of the body of bats (Gummer and Brigham, 1995; Robaina et al., 2017), turtles (Rivera and Neely, 2020), and birds (Aparicio and Bonal, 2002). Our results of decoupled patterns of longitudinal and cross-sectional humeral FA point toward independent trajectories of bone elongation (endochondral ossification) and thickening (intramembranous ossification) while also suggesting that it could be applicable within single structures. Longitudinal and cross-sectional growth are hypothesized to be decoupled ontogenetic processes (Enlow, 1963; Montoya-Sanhueza et al., 2019) and to respond to different selective pressures, with cross-sectional bone deposition potentially associated with biomechanical resistance against torsional and bending stresses (Blackburn, 2011; Perchalski et al., 2018) and bone elongation correlating with maintenance of body proportions within a functional unit (Ueti et al., 2015).

Peaks of longitudinal (stages 2 and 6) and cross-sectional FA (stage 5) at different developmental stages could indicate variations in the timing of ossification onset and/or growth rates across species, following the morphogenetic drift model (Hallgrímsson, 1998). Despite bats showing a general mammalian developmental pattern (López-Aguirre et al., 2019a), heterochronies and diverging allometric trajectories for the ossification of the humeri have been found across bat species (López-Aguirre et al., 2019b). The reported slower ossification of the humerus in yinpterochiropteran bats could indicate that different growth rates at a given developmental stage could result in increased magnitudes of FA (López-Aguirre et al., 2019b). However, our study does not allow for the accurate comparison of growth rates across species, reflecting the challenge of amassing embryonic material for non-model species, as not all species have complete developmental series. Future studies should focus on further exploring the relationship between interspecific variation in growth rates and magnitudes of FA. Despite both longitudinal and cross-sectional FA showing homogenous magnitudes across development, only cross-sectional FA showed statistically significant differences across developmental stages. Increasing postcranial morphological disparity and integration across prenatal development in bats has been reported in ossification sequences and metric growth (López-Aguirre et al., 2019a).

We hypothesize that our results (significant differences in cross-sectional FA across stages) indicate a greater selective pressure to canalize cross-sectional asymmetry as a response to functional demands associated with a newborn bat's ecology. In particular, there may be a greater requirement for symmetrical cross-sectional growth to facilitate early roosting behaviors. In many bat species, newborns attach to the mothers using their feet and thumb (Koyabu and Son, 2014) rather than to immediately fly, with bone elongation continuing during this period. Multiple studies have documented postnatal development of flight in bats, describing the altricial development of the forelimb in bats followed by accelerated bone elongation (Hughes and Rayner, 1993; Kunz and Robson, 1995; Kunz et al., 2009; Lin et al., 2010; Eghbali et al., 2017; Eghbali and Sharifi, 2018). First flights in most species that have been studied occur consistently in synchrony with weaning, usually a couple of weeks after birth once adult body dimensions are reached (Hughes and Rayner, 1993; Kunz and Robson, 1995; Kunz et al., 2009; Lin et al., 2010; Eghbali et al., 2017; Eghbali and Sharifi, 2018). We hypothesize that because self-powered flight is not achieved immediately after birth, bone elongation asymmetry and compensatory growth to optimize wing proportions would be less constrained prenatally (Ueti et al., 2015). Further studies quantifying the developmental trajectories of asymmetry should focus on describing the complementary developmental process (pre- and postnatal). Additionally, testing this hypothesis in mammal species with newborns that also cling to the mothers (e.g., macaques and colugos) would further elucidate whether the patterns found in this study are common across Mammalia.

We found significant support for the presence of FA and DA during the prenatal development of the humerus in bats. We also show that magnitudes of FA remain stable across prenatal development, and we hypothesize this to be evidence of developmental control of FA. Moreover, we find evidence for decoupled patterns of longitudinal and cross-sectional asymmetry throughout prenatal humeral development. We suggest that decoupled morphogenetic processes (i.e., bone thickening *via* intramembranous ossification and bone elongation *via* endochondral ossification) and functional differences between bone elongation and cross-sectional bone deposition may be associated with the newborn's ecology (i.e., pup roosting behavior and the later acquisition of flight). To our knowledge, this study is the first to analyze asymmetry patterns in the development of bats, providing new information about phenotypic asymmetry and DI in non-model taxa. We highlight the importance of assessing the correlation between FA and DI beyond patterns of total asymmetry FA and DI.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institute of Ecology and Biological Resources, Vietnam Academy of Sciences ethics permit No. 322/STTNSV.

AUTHOR CONTRIBUTIONS

CL-A, SH, and LW designed the study. CL-A gathered and analyzed the data. CL-A, SH, LW, and DK wrote the manuscript. DK and VTT sampled the specimens. DK prepared and scanned the specimens. CL-A processed the scans. All authors read and approved the final manuscript.

FUNDING

CL-A was supported by the Research Training Program scholarship from the Australian Department of Education, City University of Hong Kong Start-up Grant #9610466 and JSPS #18H04816, #18H02492, #18K19359, and #18KK0207, and JRPs-LEAD with DFG to DK. LW was supported by the Australian Research Council (FT200100822), and SH by Australian Research Council (DP180100792).

ACKNOWLEDGMENTS

We acknowledge the generous assistance of the directorates of the Vietnam National Parks and Nature Reserves where DK and VTT conducted the surveys and Vietnam Administration of Forestry of the Vietnamese Ministry of Agriculture and Rural Development and People's Committees of Provinces.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.639522/full#supplementary-material

REFERENCES

Adams, D. C., Otárola-Castillo, E., and Paradis, E. (2013). Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* 4, 393–399. doi: 10.1111/2041-210x.12035

Adams, R. A. (1992). Comparative skeletogenesis of the forearm of the little brown bat (*Myotis-Lucifugus*) and the Norway Rat (*Rattus-Norvegicus*). *J. Morphol.* 214, 251–260. doi: 10.1002/jmor.1052140302

Adams, R. A. (2008). Morphogenesis in bat wings: linking development, evolution and ecology. *Cells Tissues Organs* 187, 13–23. doi: 10.1159/000109960

Adams, R. A., and Shaw, J. B. (2013). "Time's arrow in the evolutionary development of bat flight," in *Bat Evolution, Ecology, and Conservation*, eds R. Adams and S. Pedersen (New York, NY: Springer), 21–46. doi: 10.1007/978-1-4614-7397-8_2

Allard, P., and Tabin, C. J. (2009). Achieving bilateral symmetry during vertebrate limb development. *Semin. Cell Dev. Biol.* 20, 479–484. doi: 10.1016/j.semcdb.2008.10.011

Aparicio, J. (1998). Patterns of fluctuating asymmetry in developing primary feathers: a test of the compensational growth hypothesis. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 265, 2353–2357. doi: 10.1098/rspb.1998.0583

Aparicio, J. M., and Bonal, R. (2002). Why do some traits show higher fluctuating asymmetry than others? A test of hypotheses with tail feathers of birds. *Heredity* 89, 139–144. doi: 10.1038/sj.hdy.6800118

Atchley, W. R., and Hall, B. K. (1991). A model for development and evolution of complex morphological structures. *Biol. Rev.* 66, 101–157. doi: 10.1111/j.1469-185X.1991.tb01138.x

Blackburn, A. (2011). Bilateral asymmetry of the humerus during growth and development. *Am. J. Phys. Anthropol.* 145, 639–646. doi: 10.1002/ajpa.21555

Broek, C. M. A. T., Bots, J., Bugiani, M., Galis, F., and Dongen, S. V. (2017). Developmental origins of limb developmental instability in human fetuses: many abnormalities make the difference. *Symmetry* 9:51. doi: 10.3390/sym9040051

Brown, E. E., Cashmore, D. D., Simmons, N. B., and Butler, R. J. (2019). Quantifying the completeness of the bat fossil record. *Palaeontology* 62, 757–776. doi: 10.1111/pala.12426

Camacho, J., Heyde, A., Bhullar, B-A. S., Haelewaters, D., Simmons, N. B., and Abzhanov, A. (2019). Peramorphosis, an evolutionary developmental mechanism in neotropical bat skull diversity. *Dev. Dyn.* 248, 1129–1143. doi: 10.1002/dvdy.90

Castillo-Figueroa, D. (2018). Fluctuating asymmetry of three bat species in extensive livestock systems from Córdoba department, Colombia. *Rev. Colomb. Ciencia Animal Recia* 10, 143–153. doi: 10.24188/recia.v10.n2.2018.623

Cooper, L., Cretekos, C. J., and Sears, K. E. (2012). The evolution and development of mammalian flight. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 773–779. doi: 10.1002/wdev.50

Cretekos, C. J., Wang, Y., Green, E. D., Martin, J. F., Rasweiler, J. J. T., and Behringer, R. R. (2008). Regulatory divergence modifies limb length between mammals. *Genes Dev.* 22, 141–151. doi: 10.1101/gad.1620408

Cretekos, C. J., Weatherbee, S. D., Chen, C. H., Badwaik, N. K., Niswander, L., Behringer, R. R., et al. (2005). Embryonic staging system for the short-tailed fruit bat, *Carollia perspicillata*, a model organism for the mammalian order Chiroptera, based upon timed pregnancies in captive-bred animals. *Dev. Dyn.* 233, 721–738. doi: 10.1002/dvdy.20400

de Figueiredo, D., Gonçalves Miller, B., Barbosa Leal, E. S., and Montes, M. A. (2015). Fluctuating asymmetry in populations of bats: species adapted to urban environments are not hampered by habitat degradation. *Chiropt. Neotrop.* 21, 1305–1311.

Dongen, S. V. (2006). Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *J. Evol. Biol.* 19, 1727–1743. doi: 10.1111/j.1420-9101.2006.01175.x

Dumont, E. R. (2010). Bone density and the lightweight skeletons of birds. *Proc. R. Soc. B Biol. Sci.* 277, 2193–2198. doi: 10.1098/rspb.2010.0117

Eghbali, H., Shahabi, S., Najafi, N., Mehdizadeh, R., Yousefi, S., and Sharifi, M. (2017). Postnatal growth, wing development and age estimations in the Mediterranean horseshoe bat *Rhinolophus* euryale (Chiroptera: Rhinolophidae) in Kerend cave, western Iran. *Mammalia* 82, 276–287. doi: 10.1515/mammalia-2017-0006

Eghbali, H., and Sharifi, M. (2018). Postnatal growth, age estimation, and wing development in Geoffroy's Bat *Myotis emarginatus* (Chiroptera: Vespertilionidae). *Mammal Study* 43, 153–165. doi: 10.3106/ms2017-0077

Enlow, D. H. (1963). *Principles of Bone Remodeling: An Account of Post-natal Growth and Remodeling Processes in Long Bones and the Mandible*. Springfield, IL: Thomas.

Farnum, C. E., Tinsley, M., and Hermanson, J. W. (2008). Forelimb versus Hindlimb Skeletal development in the big brown bat, *Eptesicus fuscus:* functional divergence is reflected in chondrocytic performance in autopodial growth plates. *Cells Tissues Organs* 187, 35–47. doi: 10.1159/000109962

Gummer, D. L., and Brigham, R. M. (1995). Does fluctuating asymmetry reflect the importance of traits in little brown bats (*Myotis lucifugus*)? *Can. J. Zool.* 73, 990–992. doi: 10.1139/z95-116

Hallgrímsson, B. (1998). Fluctuating asymmetry in the mammalian skeleton: evolutionary and developmental implications. *Evol. Biol.* 30, 187–251. doi: 10.1007/978-1-4899-1751-5_6

Hallgrímsson, B. (1999). Ontogenetic patterning of skeletal fluctuating asymmetry in rhesus macaques and humans: evolutionary and developmental implications. *Int. J. Primatol.* 20, 121–151. doi: 10.1023/A:1020540418554

Hallgrímsson, B., Miyake, T., Wilmore, K., and Hall, B. K. (2003). Embryological origins of developmental stability: size, shape and fluctuating asymmetry in prenatal random bred mice. *J. Exp. Zool. Part B Mol. Dev. Evol.* 296, 40–57. doi: 10.1002/jez.b.15

Hughes, P., and Rayner, J. M. V. (1993). The flight of pipistrelle bats *Pipistrellus pipistrellus* during pregnancy and lactation. *J. Zool.* 230, 541–555. doi: 10.1111/j.1469-7998.1993.tb02705.x

Juste, J., López-González, C., and Strauss, R. E. (2001a). Analysis of asymmetries in the African fruit bats Eidolon helvum and *Rousettus egyptiacus* (Mammalia: Megachiroptera) from the islands of the Gulf of Guinea. I. Variance and size components of bilateral variation. *J. Evol. Biol.* 14, 663–671. doi: 10.1046/j.1420-9101.2001.00298.x

Juste, J., López-González, C., and Strauss, R. E. (2001b). Analysis of asymmetries in the African fruit bats Eidolon helvum and *Rousettus egyptiacus* (Mammalia: Megachiroptera) from the islands of the Gulf of Guinea. II. Integration and levels of multivariate fluctuating asymmetry across a geographical range. *J. Evol. Biol.* 14, 672–680. doi: 10.1046/j.1420-9101.2001.00299.x

Kellner, J. R., and Alford, R. A. (2003). The ontogeny of fluctuating asymmetry. *Am. Natural.* 161, 931–947. doi: 10.1086/375177

Klingenberg, C. P. (2015). Analyzing fluctuating asymmetry with geometric morphometrics: concepts, methods, and applications. *Symmetry* 7, 843–934. doi: 10.3390/sym7020843

Klingenberg, C. P., and McIntyre, G. S. (1998). Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with procrustes methods. *Evolution* 52, 1363–1375. doi: 10.1111/j.1558-5646.1998.tb02018.x

Klingenberg, C. P., and Nijhout, H. F. (1999). Genetics of fluctuating asymmetry: a developmental model of developmental instability. *Evolution* 53, 358–375. doi: 10.1111/j.1558-5646.1999.tb03772.x

Koyabu, D., and Son, N. (2014). Patterns of postcranial ossification and sequence heterochrony in bats: life histories and developmental trade-offs. *J. Exp. Zool. Part B Mol. Dev. Evol.* 322, 607–618. doi: 10.1002/jez.b.22581

Kronenberg, H. M. (2003). Developmental regulation of the growth plate. *Nature* 423, 332–336. doi: 10.1038/nature01657

Kunz, T. H., Adams, R. A., and Wood, W. R. (2009). "Methods for assessing size at birth and postnatal growth and development in bats," in *Ecological and behavioral methods for the study of bats*, eds T. H. Kunz and S. Parsons (Baltimore, MD: The Johns Hopkins University Press), 273–314.

Kunz, T. H., and Robson, S. K. (1995). Postnatal growth and development in the mexican free-tailed bat (*Tadarida brasiliensis* mexicana): birth size, growth rates, and age estimation. *J. Mammal.* 76, 769–783. doi: 10.2307/1382746

Leamy, L. J., and Klingenberg, C. P. (2005). The genetics and evolution of fluctuating asymmetry. *Annu. Rev. Ecol. Evol. Syst.* 36, 1–21. doi: 10.1146/annurev.ecolsys.36.102003.152640

Lin, A., Jin, L., Liu, Y., Sun, K., and Feng, J. X. (2010). Postnatal growth and age estimation in Horsfield's Leaf-Nosed Bat *Hipposideros larvatus*. *Zool. Stud.* 49, 789–796.

López-Aguirre, C., Hand, S. J., Koyabu, D., Son, N. T., and Wilson, L. A. B. (2019a). Postcranial heterochrony, modularity, integration and disparity in the prenatal ossification in bats (Chiroptera). *BMC Evol. Biol.* 19:75. doi: 10.1186/s12862-019-1396-1

López-Aguirre, C., Hand, S. J., Koyabu, D., Son, N. T., and Wilson, L. A. B. (2019b). Prenatal allometric trajectories and the developmental basis of postcranial phenotypic diversity in bats (Chiroptera). *J. Exp. Zool. Part B Mol. Dev. Evol.* 332, 36–49. doi: 10.1002/jez.b.22846

López-Aguirre, C., and Pérez-Torres, J. (2015). Asimetría cráneo-mandibular de Artibeus lituratus (Chiroptera, Phyllostomidae) en Colombia. *Univ. Scientiarum* 20, 141–152. doi: 10.11144/javeriana.sc20-1.acal

López-Aguirre, C., Wilson, L. A., Koyabu, D., Tan Tu, V., and Hand, S. J. (2019). Decoupled morphological and biomechanical evolution and diversification of the wing in bats. *EcoEvoRxiv* 31, 1–31. doi: 10.32942/osf.io/k3y5f

Lüpold, S., McElligott, A. G., and Hosken, D. J. (2004). Bat genitalia: allometry, variation and good genes. *Biol. J. Linn. Soc.* 83, 497–507. doi: 10.1111/j.1095-8312.2004.00407.x

Macintosh, A. A., Davies, T. G., Ryan, T. M., Shaw, C. N., and Stock, J. T. (2013). Periosteal versus true cross-sectional geometry: a comparison along humeral, femoral, and tibial diaphyses. *Am. J. Phys. Anthropol.* 150, 442–452. doi: 10.1002/ajpa.22218

Monteiro, L. R., Mellado, B., Nogueira, M. R., and de Morais-Jr, M. M. (2019). Individual asymmetry as a predictor of fitness in the bat *Carollia perspicillata*. *J. Evol. Biol.* 32, 1207–1229. doi: 10.1111/jeb.13522

Montoya-Sanhueza, G., Wilson, L. A. B., and Chinsamy, A. (2019). Postnatal development of the largest subterranean mammal (*Bathyergus suillus*): morphology, osteogenesis, and modularity of the appendicular skeleton. *Dev. Dyn.* 248, 1101–1128. doi: 10.1002/dvdy.81

Palmer, A. R. (1994). "Fluctuating asymmetry analyses: a primer," in *Developmental Instability: Its Origins and Evolutionary Implications: Proceedings of the International Conference on Developmental Instability: Its Origins and Evolutionary Implications, Tempe, Arizona*, ed. T. A. Markow (Dordrecht: Springer), 335–364. doi: 10.1007/978-94-011-0830-0_26

Panyutina, A., Korzun, L. P., and Kuznetsov, A. N. (2015). *Flight of Mammals: From Terrestrial Limbs to Wings*. New York, NY: Springer.

Papadimitriou, H. M., Swartz, S. M., and Kunz, T. H. (1996). Ontogenetic and anatomic variation in mineralization of the wing skeleton of the Mexican free-tailed bat, *Tadarida brasiliensis*. *J. Zool.* 240, 411–426. doi: 10.1111/j.1469-7998.1996.tb05295.x

Perchalski, B., Placke, A., Sukhdeo, S. M., Shaw, C. N., Gosman, J. H., Raichlen, D. A., et al. (2018). Asymmetry in the cortical and trabecular bone of the human humerus during development. *Anat. Rec. (Hoboken)* 301, 1012–1025. doi: 10.1002/ar.23705

Rivera, G., and Neely, C. M. D. (2020). Patterns of fluctuating asymmetry in the limbs of freshwater turtles: are more functionally important limbs more symmetrical? *Evolution* 74, 660–670. doi: 10.1111/evo.13933

Robaina, Z., Denis, D., and Sánchez-Lozada, M. (2017). ¿Son regulados los niveles de asimetría fluctuante según la función de las estructuras morfológicas? Estudio de caso en Phyllonycteris poeyi (chiroptera: phyllostomidae)/Are levels of fluctuating asymmetry regulated according to the function of. *Poeyana* 505, 15–21.

Sears, K. E., Behringer, R. R., Rasweiler, J. J., and Niswander, L. A. (2006). Development of bat flight: morphologic and molecular evolution of bat wing digits. *Proc. Natl. Acad. Sci. U.S. A.* 103, 6581–6586. doi: 10.1073/pnas.0509716103

Spoor, C. F., Zonneveld, F. W., and Macho, G. A. (1993). Linear measurements of cortical bone and dental enamel by computed-tomography - applications and problems. *Am. J. Phys. Anthropol.* 91, 469–484. doi: 10.1002/ajpa.1330910405

Swaddle, J. P., and Witter, M. S. (1997). On the ontogeny of developmental stability in a stabilized trait. *Proc. R. Soc. Lond. B Biol. Sci.* 264, 329–334.

Swartz, S. M. (1997). Allometric patterning in the limb skeleton of bats: implications for the mechanics and energetics of powered flight. *J. Morphol.* 234, 277–294.

Swartz, S. M., Bennett, M. B., and Carrier, D. R. (1992). Wing bone stresses in free flying bats and the evolution of skeletal design for flight. *Nature* 359, 726–729.

Tocts, A. M. S., Johnson, D. W., and Carter, A. J. R. (2016). Strong nonlinear selection against fluctuating asymmetry in wild populations of a marine fish. *Evolution* 70, 2899–2908. doi: 10.1111/evo.13092

Ueti, A., Pompeu, P. S., and Ferreira, R. L. (2015). Asymmetry compensation in a small vampire bat population in a cave: a case study in Brazil. *Subterranean Biol.* 15, 57–67. doi: 10.3897/subtbiol.15.4807

Voigt, C. C., Heckel, G., and Mayer, F. (2005). Sexual selection favours small and symmetric males in the polygynous greater sac-winged bat *Saccopteryx bilineata* (Emballonuridae, Chiroptera). *Behav. Ecol. Sociobiol.* 57, 457–464. doi: 10.1007/s00265-004-0874-6

Webber, T., Patel, S. P., Pensak, M., Fajolu, O., Rozental, T. D., and Wolf, J. M. (2015). Correlation between distal radial cortical thickness and bone mineral density. *J. Hand Surg.* 40, 493–499. doi: 10.1016/j.jhsa.2014.12.015

Wilson, L., and Humphrey, L. T. (2015). A Virtual geometric morphometric approach to the quantification of long bone bilateral asymmetry and cross-sectional shape. *Am. J. Phys. Anthropol.* 158, 541–556. doi: 10.1002/ajpa.22809

Wolpert, L. (2010). Arms and the man: the problem of symmetric growth. *PLoS Biol.* 8:e1000477. doi: 10.1371/journal.pbio.1000477

Conflict of Interest: 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 © 2021 López-Aguirre, Hand, Koyabu, Tu and Wilson. 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) and the copyright owner(s) 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 Ecology and Evolution

PERSPECTIVE
published: 04 June 2021
doi: 10.3389/fevo.2021.666722

Innovation Through Heterochrony: An Amphioxus Perspective on Telencephalon Origin and Function

*Thurston Lacalli**

Department of Biology, University of Victoria, Victoria, BC, Canada

Heterochrony has played a key role in the evolution of invertebrate larval types, producing "head larvae" in diverse taxa, where anterior structures are accelerated and specialized at the expense of more caudal ones. For chordates, judging from amphioxus, the pattern has been more one of repeated acceleration of adult features so that they function earlier in development, thus converting the ancestral larva, whether it was a head larva or not, into something progressively more chordate-like. Recent molecular data on gene expression patterns in the anterior nerve cord of amphioxus point to a similar process being involved in the origin of the telencephalon. As vertebrates evolved, a combination of acceleration and increasing egg size appears here to have allowed the development of a structure that would originally have emerged only gradually in the post-embryonic phase of the life history to be compressed into embryogenesis. The question then is what, in functional terms, makes the telencephalon so important to the survival of post-embryonic ancestral vertebrates that this was adaptively advantageous. A better understanding of the function this brain region performs in amphioxus may help provide the answer.

Keywords: developmental timing, head larvae, forebrain prosomeres, olfaction, cortical neurons

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University of Málaga, Spain

Reviewed by:
Hector Escriva,
Centre National de la Recherche Scientifique (CNRS), France
Sebastian Shimeld,
University of Oxford, United Kingdom

****Correspondence:***
Thurston Lacalli
lacalli@uvic.ca

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Ecology and Evolution

Received: *16 February 2021*
Accepted: *29 April 2021*
Published: *04 June 2021*

Citation:
Lacalli T (2021) Innovation Through Heterochrony: An Amphioxus Perspective on Telencephalon Origin and Function. Front. Ecol. Evol. 9:666722. doi: 10.3389/fevo.2021.666722

INTRODUCTION

One important way that development contributes to evolutionary innovation is through heterochrony, the process of accelerating or delaying some developmental events with respect to others (de Beer, 1971). This is most dramatic when it results in the emergence of a novel life history stage, as in the "head larvae" of various marine invertebrate taxa, where head structures are accelerated at the expense of more caudal ones, and become secondarily specialized to suit the larval mode of life. This accounts for features ranging from the suppressed trunk of many spiralians to the dipleurula-type larvae of echinoderms and hemichordates (Lacalli, 2005; Gonzalez et al., 2017), but the degree to which such changes are basal within bilaterians or are taxon-specific is a matter of debate (Strathmann, 2020). There also is a countertrend in larval specialization involving the acceleration of adult features so they function in the larva, a process known as adultation, which has occurred independently in multiple taxa (Jagersten, 1972).

How these ideas apply to chordates is not entirely clear, as chordate ancestry may never have included a true head larva like the dipleurula. But heterochrony would explain how locomotory specializations originally adapted for adult life (somites, nerve cord and notochord in the case of chordates) would progressively convert the ancestral larva, whatever it was like, into something anatomically more chordate-like. And, though newly hatched amphioxus have been justifiably referred to as "swimming heads" (Gilland and Baker, 1993), this is because development proceeds in an antero-posterior sequence, and only a limited number of anterior somites can be produced from the nutrient supply provided by the egg. With this framework in mind, the purpose of this brief account is to explore the implications of new data on another example of heterochrony involving what is evidently a chordate-specific neuroanatomical innovation, the telencephalon.

TELENCEPHALON ORIGINS AND ANCESTRAL FUNCTION

Most molecular data on amphioxus neural organization depend on studies on the early larva (reviewed by Holland and Holland, 2021). These reveal little about any putative telencephalon homolog because the antero-dorsal region from which it would necessarily form is largely undifferentiated at that time (Wicht and Lacalli, 2005). The question of whether amphioxus has such a homolog may now have been resolved by the observations of Benito-Gutierrez et al. (2021), who have identified markers associated with the telencephalon in a dorsal, pre-infundibular domain (termed the pars anterodorsalis, or PAD) that is insignificant in young larvae, but expands as the larvae mature, differentiating fully only in the adult. The molecular data reveal populations of glutamatergic and GABAergic neurons, both of which occur widely in the telencephalon, and of dopaminergic neurons, found in the vertebrate olfactory bulb and hypothalamus (Yamamoto and Vernier, 2011). There are similarities also with the telencephalon in the shape of the central canal and the pattern of thickening of the neural epithelium. The degree of homology is open to question, however, and the authors are suitably cautious in their interpretation. This is appropriate because the pre-infundibular neural domain, taken as a whole, maps to the apical plate of diverse invertebrate larvae (Tosches and Arendt, 2013; Marlow et al., 2014) and hence, predates chordates, which would account for the involvement of *FoxG*, a marker for apically-derived sensory derivatives such as photoreceptor organs (Toresson et al., 1998; Aldea et al., 2015; Liu and Satou, 2019). More specific to telencephalon are *Emx* and *Nkx2.1*, and *Pax* genes (Lichtneckert and Reichert, 2005; Molyneaux et al., 2007), but here too the expression could be as indicative of origin from a common domain as it is of structure-specific homology. A further problem is to determine whether the domain belongs to the dorsal part of the corresponding vertebrate prosomere or is an alar-plate hypothalamic derivative, which, based on current ideas about forebrain patterning (Puelles and Rubenstein, 2015; Yamamoto et al., 2017) is a possibility.

A crucial but missing piece of the puzzle is the function this rostral- and dorsal-most part of the brain performs in amphioxus. If it proved to share functions with the telencephalon, that would further support the case for homology, and provide some guidance regarding the ancestral function of this brain region in the last common ancestor of cephalochordates and vertebrates. Given the positioning of the olfactory tracts and their targets in vertebrates, in rostral forebrain (Kerman et al., 2013; Suryanarayana et al., 2021), the processing of olfactory inputs is very likely to have been a core function of this part of the ancestral brain. It is then suggestive that amphioxus has a late-developing pair of rostral nerve branches that project to (but also through) the PAD (Lacalli, 2002). Nothing is known of the sensory cells supplying these nerves beyond their location, at or near the tip of the rostrum (Poncelet and Shimeld, 2020), but this is precisely where amphioxus homologs of vertebrate olfactory neurons would be expected. And, the amphioxus genome does contain sequences for vertebrate-type olfactory receptors (Churcher and Taylor, 2009). Given the apparent underdevelopment of the PAD in all stages but mature adults, one could postulate a role, presumably olfactory, in the control of reproduction and spawning behavior. But there is as yet no evidence for this or any other olfactory involvement in the control of amphioxus behavior, a reminder of how neglected such aspects of amphioxus biology have been in recent decades.

THE LIFE HISTORY CONTEXT

On the assumption Benito-Gutierrez et al. (2021) have indeed discovered an amphioxus homolog of the telencephalon, or at least a part of it, the life history context can be understood as shown in **Figure 1**: that in amphioxus, this brain region only begins its development once the pelagic feeding phase reaches sufficient size, and may only function in the mature adult. For vertebrates, in contrast, a combination of increased yolk supply and developmental acceleration has compressed the developmental sequence so it can be completed within the period of embryogenesis. This implies that having a functioning telencephalon at or shortly after hatching was of such benefit to basal vertebrates that the diversion of nutrient required to achieve this was justified by increased hatchling survival. The eyes fit this same pattern, as they are disproportionately large in the young stages of many vertebrates, including, for example, many fish larvae. The advantage of precocious development to full size is understandable in the case of the eye, as a large optic cup fitted with a lens has clear advantages in terms of resolution and image-formation over a small cluster of photoreceptors, however, shaped and positioned. In consequence, if the ancestral condition was one where a simple photoreceptor array developed progressively as the animal grew, adding elements and reshaping itself as size permitted, there should be strong selective pressure to shift the point where an image of optimal resolution could be produced to the earliest possible stage in the life history.

The same argument could apply to processing olfactory input. Olfaction is crucial for the detection of predators and kin in fish larvae of various types, a function that involves olfactory imprinting (Chivers and Smith, 1994; Gerlach and Wullimann, 2021). Whether this is considered a rudimentary form of learning or more akin to memory, it implies integrative functions of some complexity requiring a dedicated circuitry and a more than minimal tissue mass. Hence, like the eye, there would be some justification for supposing that precocious differentiation, rather

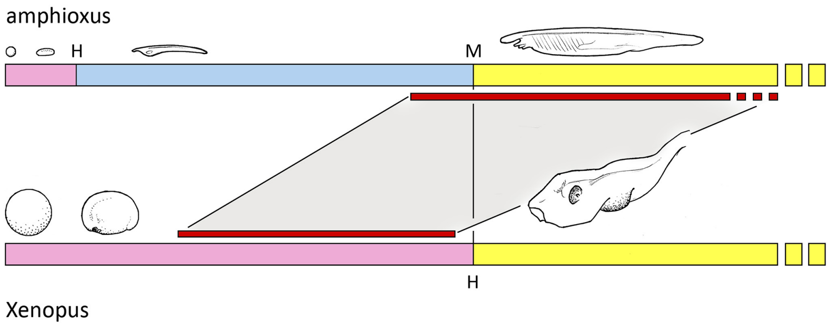


FIGURE 1 | The early life history of amphioxus compared with a representative aquatic vertebrate, *Xenopus* in this case. The embryonic stage (in pink) produces a hatching larva (at H), followed by a pelagic feeding phase (blue) that metamorphoses (M) to a juvenile that is essentially a small adult with adult habits (yellow). Embryogenesis in *Xenopus* also produces a larva, but is here shown in yellow because it corresponds less with the amphioxus larva than the juvenile in size and anatomical complexity, and the term larva is problematic anyway if, as here, the stages in question likely have independent evolutionary origins. Two trends are evident. First, that vertebrate eggs and embryos are orders of magnitude larger in volume, and so can be far more fully developed at hatching, and second, that this allows features that can only develop progressively in amphioxus to be incorporated into embryogenesis in vertebrates. The focus in this account is the telencephalon and its putative amphioxus homolog, whose periods of development, inferred for amphioxus from data reported by Benito-Gutierrez et al. (2021), are shown by the red bars.

than a gradual increase in size and cell numbers, could have been the better option from an adaptive standpoint. Conversely, for early vertebrates, a fully functional telencephalon may have been needed, even in young hatchlings, to meet the demands of an emerging predatory mode of life, e.g., to coordinate the visual tracking and capture of prey. Amphioxus would then be important primarily as a benchmark, defining a starting point for these and other subsequent changes.

NEURONAL CELL TYPES AND THEIR ORIGINS

It is worth considering also the evolutionary origin of the cell types identified by Benito-Gutierrez et al. (2021), especially the glutamatergic and GABAergic neurons which by implication, would be related by homology to the main excitatory and inhibitory cell classes, respectively, in vertebrate cortex. Two categories of dorsal post-infundibular amphioxus neurons, the translumenal and dorsal bipolar neurons, are noteworthy here for being apico-basally inverted (Lacalli and Kelly, 2003), a feature also characteristic of cortical neurons (Barnes and Polleux, 2009). However, both types of amphioxus neurons appear to be absent from the pre-infundibular region of the adult nerve cord (Castro et al., 2015), which may rule out a direct relation, through homology, with particular classes of telencephalic neurons. It would be useful to know whether the glutamatergic neurons identified by Benito-Gutierrez et al. (2021) will, on further investigation, show signs of inversion, or whether this feature is, for the telencephalon at least, a vertebrate innovation. In either case, and despite similar transmitters being involved, it is entirely possible that the dorsal neurons populating the cerebral cortex are only distantly related in evolutionary terms to those in midbrain and hindbrain, i.e., tectum and cerebellum, and possess unique functional capabilities that account for the division of labor between these structures. Molecular criteria for distinguishing among the relevant cell types are only beginning to be applied across vertebrates (e.g., Tosches and Laurant, 2019), but in future could be usefully extended to include amphioxus.

DISCUSSION

The results of Benito-Gutierrez et al. (2021) can be interpreted as indicting strongly that the telencephalon, as a structure with particular functions, is a comparatively late innovation that is chordate-specific in its elaboration. Understanding the sequence of events that led to its emergence as a major brain region in vertebrates then depends on knowing its function in basal chordates. A role in the processing of olfactory input probably predates other functions, but there could be additional tasks preformed by this brain region, also traceable back to the common ancestor of amphioxus and vertebrates, of which we are as yet unaware. To resolve this issue, more research on adult amphioxus will be required, both on behavior and sensory physiology, but especially on neuroanatomy at the cellular level, directed toward compiling an inventory of the neuronal cell types in the rostral brain, their morphology and synaptic connections. Ideally one would like to have a complete map of synaptic patterns, i.e., the connectome, for this brain region. This has already been done in large part for young larvae, and adult stages would not present much more of a technical challenge given recent improvements in methodology (Landhuis, 2020). Compared to the brain of a mouse or larval zebrafish, an analysis of the amphioxus brain is orders of magnitude less demanding, with the promise of meaningful results in a correspondingly shorter time.

DATA AVAILABILITY STATEMENT

There is no data beyond that included in the article; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms he is the sole contributor to this work and has approved it for publication.

FUNDING

Funds to support this work were received from the L. G. Harrison Research Trust.

ACKNOWLEDGMENTS

I thank Nick and Linda Holland for their comments, Detlev Arendt for providing information on the amphioxus PAD before publication, and Riley Lacalli for assisting with the figure. I take this opportunity also to acknowledge R. J. F. (Jan) Smith, who introduced me to the alarming world of fish behavior. His untimely death cut short plans for a joint, collaborative investigation of amphioxus olfaction.

REFERENCES

Aldea, D., Leon, A., Bertrand, S., and Escriva, H. (2015). Expression of Fox genes in the cephalochordate *Branchiostoma lanceolatum*. *Front. Ecol. Evol.* 3:80. doi: 10.3389/fevo.2015.00080

Barnes, A. P., and Polleux, F. (2009). Establishment of axon-dendrite polarity in developing neurons. *Annu. Rev. Neurosci.* 32, 347–381. doi: 10.1146/annurev.neuro.31/060407.125536

Benito-Gutierrez, E., Gattoni, G., Stemmer, M., Rohr, S. D., Schuhmacher, L. N., Tang, J., et al. (2021). The dorsoanterior brain of adult amphioxus shares similarities in expression profile and neuronal composition with the vertebrate telencephalon. *BMC Biol.* 19:110. doi: 10.1186/s12915-021-01045-w

Castro, A., Becerra, M., Manso, M. J., and Anadon, R. (2015). Neuronal organization in the brain in the adult amphioxus (*Branchiostoma lanceolatum*): a study with acetylated tubulin immunohistochemistry. *J. Comp. Neurol.* 523, 2211–2232. doi: 10.1002/cne.23785

Chivers, D. P., and Smith, R. J. F. (1994). The role of experience and chemical signaling in predator recognition by fathead minnows. *J. Fish. Biol.* 44, 273–285. doi: 10.1111/j.1095-8649.1994.tb01205.x

Churcher, A. M., and Taylor, J. S. (2009). Amphioxus (*Branchiostoma floridae*) has orthologs of vertebrate odorant receptors. *BMC Evol. Biol.* 9:242. doi: 10.1186/1471-2148-9-242

de Beer, G. (1971). *Embryos and Ancestors*. Oxford: Clarendon Press.

Gerlach, G., and Wullimann, M. F. (2021). Neural pathways of olfactory kin imprinting and kin recognition in zebrafish. *Cell Tiss. Res.* 383, 273–287. doi: 10.1007/s00441-020-03378-4

Gilland, E., and Baker, R. (1993). Conservation of neuroepithelial and mesodermal segments in the embryonic vertebrate head. *Acta Anat.* 148, 110–123. doi: 10.1159/000147530

Gonzalez, P., Uhlinger, K. R., and Lowe, C. J. (2017). The adult body plan of indirect developing hemichordates develops by adding a Hox-patterned trunk to an anterior larval territory. *Curr. Biol.* 27, 87–95. doi: 10.1016/j.cub.2016.10.047

Holland, L. Z., and Holland, N. D. (2021). Cephalochordates: a window into vertebrate origins. Curr. Top. Dev. Biol. 141, 119–147. doi: 10.1016/bs.ctdb.2020.07.001

Jagersten, G. (1972). *Evolution of the Metazoan Life Cycle: A Comprehensive Theory*. New York, NY: Academic Press.

Kerman, F., Franco, L. M., Wyatt, L., and Yaksi, E. (2013). Neural circuits mediating olfactory-driven behavior in fish. *Front. Neural Circuits* 7:62. doi: 10.3389/fncr.2013.00062

Lacalli, T. C. (2002). Sensory pathways in amphioxus larvae I. constituent fibres of the rostral and anterodorsal nerves, their targets and evolutionary significance. *Acta Zool.* 83, 149–166. doi: 10.1046/j.1463-6395.2003.00109.x

Lacalli, T. C. (2005). Protochordate body plan and the evolutionary role of larvae: old controversies resolved? *Can. J. Zool.* 83, 216–224. doi: 10.1139/Z04-162

Lacalli, T. C., and Kelly, S. J. (2003). Sensory pathways in amphioxus larvae II. dorsal tracts and translumenal cells. *Acta Zool.* 84, 1–13. doi: 10.1046/j.1463-6395.2002.00065.x

Landhuis, E. (2020). Probing fine-scale connections in the brain. *Nature* 586, 631–633. doi: 10.1038/d41586-020-02947-5

Lichtneckert, R., and Reichert, H. (2005). Insights into the urbilaterian brain: conserved genetic patterning mechanisms in insect and vertebrate brain development. *Heredity* 94, 465–477. doi: 10.1038/sj.hdy.6800664

Liu, B., and Satou, Y. (2019). Foxg specifies sensory neurons on the anterior neural plate border of the ascidian embryo. *Nat. Comm.* 10:4911. doi: 10.1038/s41467-019-12839-6

Marlow, H., Tosches, M. A., Tomer, R., Steinmetz, P. R., Lauri, A., and Arendt, D. (2014). Larval body patterning and apical organs are conserved in animal evolution. *BMC Biol.* 12:7. doi: 10.1186/1741.7007.12.7

Molyneaux, B. J., Arlotta, P., Menezes, J. R. L., and Macklis, J. D. (2007). Neuronal subtype specification in the cerebral cortex. *Nat. Rev. Neurosci.* 8, 427–437. doi: 10.1038/nrn2151

Poncelet, G., and Shimeld, S. M. (2020). The evolutionary origin of the vertebrate olfactory system. *Open Biol.* 10:200330. doi: 10.1098/rsob.200330

Puelles, L., and Rubenstein, J. L. R. (2015). A new scenario of hypothalamic organization: rationale of new hypotheses introduced in the updated prosomeric model. *Front. Neuroanat.* 9:27. doi: 10.3389/fnana.2015.00027

Strathmann, R. R. (2020). Multiple origins of feeding head larvae by the early Cambrian. *Can. J. Zool.* 98, 761–776. doi: 10.1139/cjz-2019-0284

Suryanarayana, S. M., Perez-Fernandez, J., Robertson, B., and Grillner, S. (2021). Olfaction in lamprey revisited – dual projections of mitral and tufted cells. *Cell Rep.* 34:108596. doi: 10.1016/j.celrep.2020.108596

Toresson, H., Matinez-Barbera, J. P., Bardsley, A., Caubit, X., and Krauss, S. (1998). Conservation of BF-1 expression in amphioxus and zebrafish suggests evolutionary ancestry of anterior cell types that contribute to vertebrate telencephalon. *Dev. Genes. Evol.* 208, 431–439. doi: 10.1007/s004270050200

Tosches, M. A., and Arendt, D. (2013). The bilaterian forebrain: an evolutionary chimera. *Curr. Opin. Neurobiol.* 23, 1–10. doi: 10.1016/j.conb.2013.09.005

Tosches, M. A., and Laurant, G. (2019). Evolution of neural identity in cerebral cortex. *Curr. Opin. Neurobiol.* 56, 199–208. doi: 10.1016/j.conb.2019.04.009

Wicht, H., and Lacalli, T. C. (2005). The nervous system of amphioxus: structure, development, and evolutionary signficance. *Can. J. Zool.* 83, 122–150. doi: 10.1139/Z04-163

Yamamoto, K., and Vernier, P. (2011). The evolution of dopamine systems in chordates. *Front. Neuroanat.* 5:21. doi: 10.3389/fnana.2011.00021

Yamamoto, K., Bloch, S., and Vernier, P. (2017). New perspective on the regionalization of the anterior forebrain in Osteichthyes. *Dev. Growth Differ.* 59, 175–187. doi: 10.1111/dgd.12348

Conflict of Interest: 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.

Copyright © 2021 Lacalli. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 09 June 2021
doi: 10.3389/fcell.2021.700827

Massive Gene Loss and Function Shuffling in Appendicularians Stretch the Boundaries of Chordate Wnt Family Evolution

Josep Martí-Solans[1†], Hector Godoy-Marín[1], Miriam Diaz-Gracia[1], Takeshi A. Onuma[2], Hiroki Nishida[2], Ricard Albalat[1*] and Cristian Cañestro[1*]

[1] Departament de Genètica, Microbiologia i Estadística, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain, [2] Department of Biological Sciences, Graduate School of Science, Osaka University, Osaka, Japan

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Shigehiro Kuraku,
RIKEN Center for Biosystems
Dynamics Research (BDR), Japan
Takehiro Kusakabe,
Konan University, Japan

***Correspondence:**
Ricard Albalat
ralbalat@ub.edu
Cristian Cañestro
canestro@ub.edu

†Present address:
Josep Martí-Solans,
Sars International Center for Marine
Molecular Biology, University
of Bergen, Bergen, Norway

Specialty section:
This article was submitted to
Evolutionary Developmental Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 26 April 2021
Accepted: 19 May 2021
Published: 09 June 2021

Citation:
Martí-Solans J, Godoy-Marín H,
Diaz-Gracia M, Onuma TA, Nishida H,
Albalat R and Cañestro C (2021)
Massive Gene Loss and Function
Shuffling in Appendicularians Stretch
the Boundaries of Chordate Wnt
Family Evolution.
Front. Cell Dev. Biol. 9:700827.
doi: 10.3389/fcell.2021.700827

Gene loss is a pervasive source of genetic variation that influences species evolvability, biodiversity and the innovation of evolutionary adaptations. To better understand the evolutionary patterns and impact of gene loss, here we investigate as a case study the evolution of the wingless (Wnt) family in the appendicularian tunicate *Oikopleura dioica*, an emergent EvoDevo model characterized by its proneness to lose genes among chordates. Genome survey and phylogenetic analyses reveal that only four of the thirteen Wnt subfamilies have survived in *O. dioica*—Wnt5, Wnt10, Wnt11, and Wnt16,—representing the minimal Wnt repertoire described in chordates. While the loss of Wnt4 and Wnt8 likely occurred in the last common ancestor of tunicates, representing therefore a synapomorphy of this subphylum, the rest of losses occurred during the evolution of appendicularians. This work provides the first complete Wnt developmental expression atlas in a tunicate and the first insights into the evolution of Wnt developmental functions in appendicularians. Our work highlights three main evolutionary patterns of gene loss: (1) conservation of ancestral Wnt expression domains not affected by gene losses; (2) function shuffling among Wnt paralogs accompanied by gene losses; and (3) extinction of Wnt expression in certain embryonic directly correlated with gene losses. Overall our work reveals that in contrast to "conservative" pattern of evolution of cephalochordates and vertebrates, *O. dioica* shows an even more radical "liberal" evolutionary pattern than that described ascidian tunicates, stretching the boundaries of the malleability of Wnt family evolution in chordates.

Keywords: gene loss, gene function shuffling, chordate evolutionary developmental biology, wingless (Wnt) family evolution, appendicularian tunicate chordate

INTRODUCTION

The bloom of genomics is providing a new vision on gene loss as one of the major sources of genetic variation with great potential to contribute to evolutionary adaptation and the generation of biodiversity, and therefore to impact on the evolvability of groups of organisms (Olson, 1999; Albalat and Cañestro, 2016; Sharma et al., 2018; Guijarro-Clarke et al., 2020; Helsen et al., 2020; Xu and Guo, 2020). Despite mutational events that lead to non-functionalization and gene loss are random, the analyses of biased patterns of gene loss can reveal relevant information about the evolutionary impact of the losses (reviewed in Albalat and Cañestro, 2016). Thus, the loss of genes in the last common ancestor of a particular taxonomical group, or differences in the trends of gene loss of certain functional categories among different taxa might reflect differences in gene essentiality/dispensability resulting from differential selective restrictions or the evolution of divergent adaptations in different taxa. The wingless (*Wnt*) family, which encodes a set of secreted glycoprotein ligands that regulate key events of animal development, is a paradigmatic example of pervasive gene loss with different trends of gene loss among different taxa (Albalat and Cañestro, 2016). While deuterostomes, for instance, conserve in general most of the *Wnt* family repertoire, many protostomes have suffered extensive losses (Prud'homme et al., 2002; Darras et al., 2018; Somorjai et al., 2018). The loss of *Wnt3* in the last common ancestor of all protostomes can be considered a synapomorphic trait that might have impacted on the origin and evolution of this clade (Cho et al., 2010).

Gene loss, however, is not always adaptive, but in many occasions occurs under neutral conditions, as a consequence, for example, of a process of regressive evolution (reviewed in Albalat and Cañestro, 2016). Increase of mutational robustness or changes of environmental conditions can lead to an increase of the dispensability of certain genes, facilitating therefore selectively neutral gene loss without significant phenotypic impact (Albalat and Cañestro, 2016) as have been recently concluded after a comprehensive comparative genomic analysis across the metazoan tree of life (Fernández and Gabaldón, 2020). Gene losses have been frequently accompanied by events of function shuffling, in which paralogs or related gene families can co-opt redundant functions, and therefore increasing mutational robustness that can favor gene loss (McClintock et al., 2001; Cañestro et al., 2009). How function shuffling occurs, however, remains unclear. Promiscuous gene families, in which the activity of different paralogs is similar and functions among genes are interchangeable, are prone to bear function shuffling. Several examples of function shuffling accompanied by gene losses have been reported among Wnt genes (Somorjai et al., 2018) and references therein), which have been related to the function promiscuity of this gene family since multiple Wnt ligands can activate more than one pathway (Ring et al., 2014). *Wnt* ligands, after binding to the Frizzled receptor, can activate two different pathways: (i) the cell-fate pathway (a.k.a. "*Wnt*/β-catenin pathway" or the "canonical *Wnt* pathway"), which is mediated by the nuclear localization of β-catenin for pathway activation; and (ii) the cell-polarity pathway, which is mediated by several intermediate effectors acting independently of β-catenin, and includes, at least, the non-canonical planar cell polarity (PCP) pathway and the non-canonical *Wnt*/Calcium pathway (Loh et al., 2016).

Despite the Wnt family plays fundamental roles during development and adult tissue homeostasis in a vastly conserved way among metazoans, it still remains unclear how this gene family has evolved so radically different patterns of gene loss in different taxa, even within the same phylum. Within the chordate phylum, for instance, recent analysis of the *Wnt* family has revealed that cephalochordates show a "conservative" pattern of evolution, retaining the complete Wnt repertoire as single-member subfamilies (Somorjai et al., 2018). Vertebrates have maintained all the Wnt subfamilies, with the exception of WntA that has been associated to the evolution of an alternative mode to open a mouth, and have expanded the number of paralogs in Wnt subfamilies through the two rounds of whole genome duplication (Somorjai et al., 2018). Within tunicates, ascidians show a "liberal" pattern of evolution including ancestral synapomorphic gene losses (i.e., Wnt4 and Wnt8), some losses affecting specific ascidian lineages (e.g., Wnt1 in *Phelobobranchia* and Wnt3 in *Molgula*), and some burst of gene duplications affecting some ascidian groups (e.g., Wnt5 in *Stolidobranchia*). This liberal pattern of evolution has been argued that might have contributed to the morphological diversification of tunicates (Somorjai et al., 2018). In appendicularian tunicates, however, the evolution of the *Wnt* family remains unknown, and considering that these organisms are highly prone to lose genes (Ferrández-Roldán et al., 2019) and other signaling pathways such as retinoic acid has been dismantled (Martí-Solans et al., 2016), it appears as an attractive system to study the impact of gene loss and the limits of *Wnt* evolution in chordates.

To address this question, we have conducted an exhaustive survey of *Wnt* genes in genomic databases of *O. dioica* and have generated the first complete atlas of developmental expression of the Wnt family in appendicularians and the first fully described in all tunicates. Our study reveals a very dynamic evolution of Wnt signaling in *O. dioica* that would have led to an extraordinary reduction of the number of subfamilies—with only 4 out of the 13 subfamilies, which represents the smallest Wnt catalog described so far in chordates—accompanied by the expansion of the Wnt11 subfamily by lineage-specific gene duplications. Our detailed atlas of *Wnt* expression in *O. dioica* reveals that expression domains encompassed tissues derived from all three germ layers in a highly dynamic manner as well as several cases of "function shuffling." Finally, our study also suggests that an asymmetrically localized maternal Wnt ligand is required for axis formation. Therefore, our results allow us to evaluate the contributions of different Wnt subfamilies during *O. dioica* development and to investigate the evolutionary and functional limits of *Wnt* signaling in chordate development.

MATERIALS AND METHODS

Laboratory Culture of *Oikopleura dioica*

O. dioica specimens and embryos were obtained from animal colonies cultured in our lab for more than 5 years and originally collected in the Mediterranean coast of Barcelona (Catalonia, Spain) as previously described (Martí-Solans et al., 2015). Ovarian microinjection was performed as previously described (Omotezako et al., 2013; **Supplementary Materials and Methods**).

Genome Database Searches and Phylogenetic Analyses

Protein sequences of the *Wnt* genes from vertebrate *Homo sapiens* and tunicate *Ciona robusta* were used as queries in

BLASTp and tBLASTn searches in *O. dioica* genome databases[1]. Homologies were assigned by phylogenetic tree analyses based on Maximum Likelihood (ML) inferences calculated with PhyML v3.0 (Guindon et al., 2010) using protein alignments generated with MUSCLE and reviewed by hand (Edgar, 2004). Robustness of tree topologies was assessed under automatic model selection based on Akaike and Bayesian Information Criteria as well as by LG, WAG, and JTT substitution models. Due to computational load limitation of bootstrap performance, branch support was inferred by fast likelihood based methods aLRT SH-like and aBayes. Accession numbers for *O. dioica* sequences are provided in **Supplementary Table 1**.

Gene Expression and Tissue Differentiation Analyses

Fragments of *O. dioica* genes were PCR amplified and cloned to synthesize gene-specific riboprobes (**Supplementary Table 2**). To reveal *Wnt* expression and evaluate neural tissues and notochord differentiation, whole-mount *in situ* hybridization on fixed embryos was performed as previously described (Bassham and Postlethwait, 2000; Cañestro and Postlethwait, 2007; Martí-Solans et al., 2016). Nuclear staining (1 μM Hoeschst in PBST for 1 h at room temperature) was included in expression analysis at tailbud stages to confirm muscle cell positions. *α-Tubulin A* and Brachyury were used as specific markers for neuronal tissues and notochord, respectively (Bassham and Postlethwait, 2000; Seo et al., 2004; Søviknes et al., 2007). Histochemical reaction of acetylcholinesterase (AChE) was used to examine the differentiation of muscle cells, while histochemical staining for alkaline phosphatase was used to monitor the differentiation of endoderm cells (Imai et al., 2000; Nishino et al., 2000; Omotezako et al., 2017). For germ-line differentiation, immunohistochemistry using an antibody against *Ciona robusta* Vasa homolog was carried out as previously reported (Onuma et al., 2017). The primary antibody used was affinity-purified rabbit anti-CiVH (1:500 dilution) (Shirae-Kurabayashi, 2006) and the secondary antibody used was Alexa Fluor 594 goat anti-rabbit IgG (1:500 dilution; Life Technologies).

RESULTS

Only Four Wnt Subfamilies Have Survived in *O. dioica*

We conducted an exhaustive survey of *Wnt* genes in genomic database of *O. dioica* that revealed the presence of 8 putative *Wnt* sequences (**Supplementary Table 1**). To classify these *Wnt* genes, we performed phylogenetic reconstructions using a total of 254 *Wnt* sequences from 20 species representing all major metazoan groups, from the cnidarian *Nematostella vectensis* to the vertebrate *Homo sapiens*. The analysis recovered all 13 *Wnt* subfamilies as monophyletic groups and distributed *O. dioica Wnt*s among these 13 subfamilies with high support values (**Figure 1**). Maximum likelihood (ML) phylogenetic tree suggested that the 8 *O. dioica Wnts* belonged to 4 Wnt subfamilies—i.e., Wnt5, Wnt10, Wnt11 (5 sequences) and Wnt16 subfamilies—(**Figure 1**). The results, therefore, show that *O. dioica* have lost 9 Wnt subfamilies during its evolution. On the other hand, our results revealed that *O. dioica* has expanded the Wnt11 subfamily to at least 4 paralogs, named *Odi_Wnt11a* to *Odi_Wnt11d*. Analysis of the gene structure of the *Odi_Wnt11* paralogs showed that *Odi_Wnt11a* had 5 introns, one of them in a conserved position in all *Wnt* genes across all metazoans (boxed black arrowhead in **Supplementary Figure 1A**), and other intron located in a position considered a signature of Wnt11 subfamily in bilaterians (Cho et al., 2010; green arrowhead in **Supplementary Figure 1**), further supporting its orthology. The other 3 *Odi_Wnt11* paralogs, namely *Odi_Wnt11b, Odi_Wnt11c*, and *Odi_Wnt11d*, had no introns, pointing to the possibility of a retrotranscriptional origin during the evolution of the appendicularian lineage. Our genome survey found a partial sequence of a potential duplicate of Wnt11c with a similarity > 94% in the genome assemblies of *O. dioica* specimens from Norway (Wnt11c_NOR2: GSOIDG00009921001) as well as from Osaka (Wnt11c_OSA2: OSKA2016.S19.g13171.01) that may have been independently duplicated in both populations (**Supplementary Figure 1B**). However, the fact that no expression of this potential Wnt11c duplicate was detected in the gene expression matrix of the Oikobase (**Supplementary Figure 6**) nor in any of the ESTs collections of any of the two populations, together with the fact that this duplicated Wnt11c could not be amplified by PCR on genomic DNA nor cDNA from specimens from Barcelona, suggested that these could be allelic variants that have been artifactually duplicated during the genome assemblies, rather than actual gene duplicates. Whole genome sequencing with a telomere-to-telomere quality from specimens of Barcelona and other locations, and further interpopulation comparisons will be needed to clarify if *Wnt11c* has been independently duplicated in different *O. dioica* populations or these sequences simply correspond to allelic variants of *Wnt11c*.

The Wnt signaling pathway, far from simple, depends on the action of multiple genes (i.e., Wnt receptors, secreted Wnt inhibitors, intermediate effectors, etc.), which we wondered if they could have been also affected by the massive loss of *Wnt* ligands in *O. dioica*. To investigate this possibility and to assess the conservation of the Wnt signaling pathways in *O. dioica*, we used the KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007) on *O. dioica* genomic and transcriptomic data from Danks et al. (2013) to automatically identify the *O. dioica* orthologs to the components of the three main Wnt signaling pathways. KAAS analysis revealed conservation of the key components of the three Wnt pathways, with the exception of Axin, APC and several antagonists (**Supplementary Figure 2A** and **Supplementary Table 1**). Phylogenetic analysis in *O. dioica* of Wnt receptor *Frizzled* (*Fzd*) (**Supplementary Figure 2B**) and Wnt antagonist *secreted Frizzled related protein* (*sFRP*) (**Supplementary Figure 2C**) revealed, in addition, a decrease in the diversity of these components in this species. *O. dioica* appeared to have lost 2 out of the 5 Fzd subfamilies present in ascidians and vertebrates –retaining members in the Fdz1/2/7 (1

---

[1] http://oikoarrays.biology.uiowa.edu/Oiko

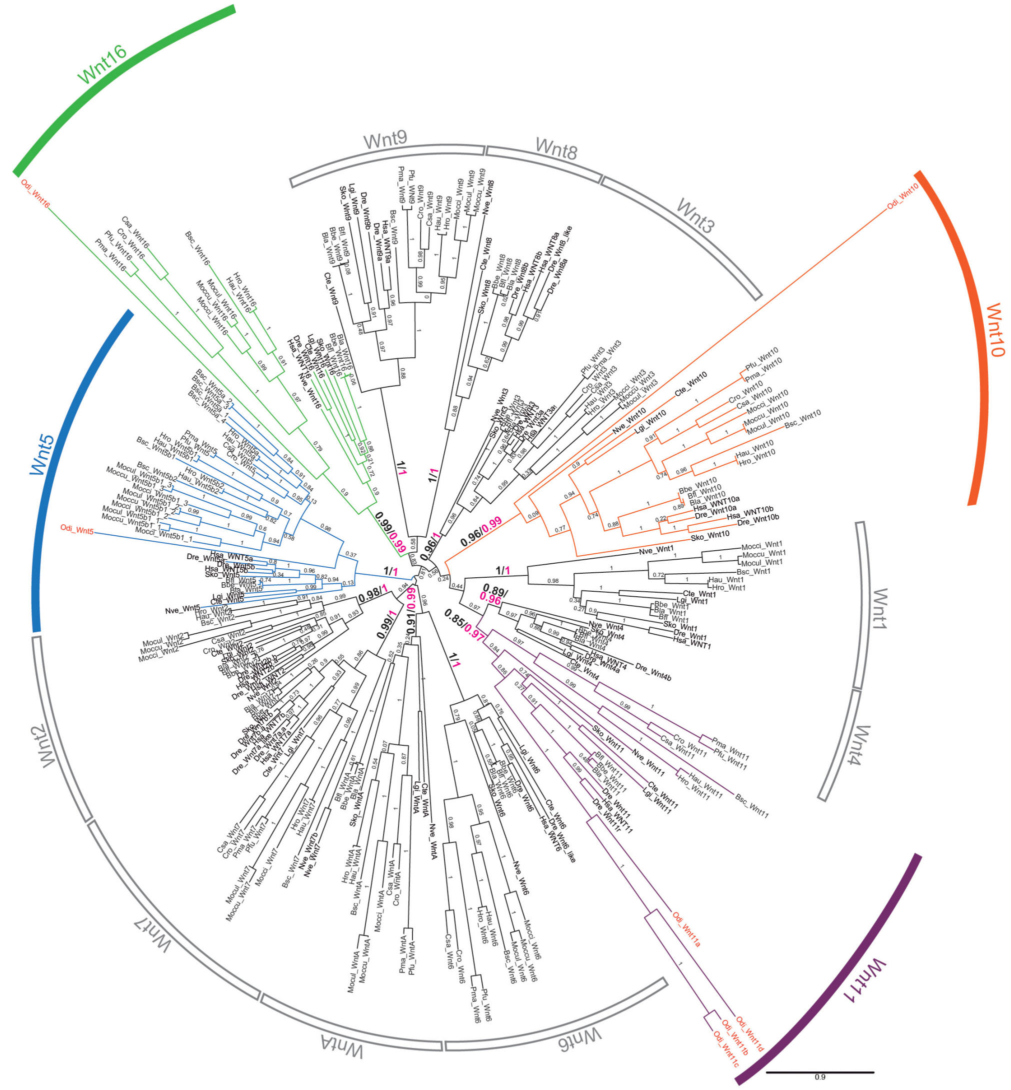


FIGURE 1 | Phylogenetic analysis of the *Wnt* family. ML phylogenetic tree of the *Wnt* family in chordates reveals that *Wnt* gene repertoire in *Oikopleura dioica* (names in red) belongs to four subfamilies: Wnt 5 (blue), Wnt10 (orange), Wnt11 (purple), and Wnt16 (green). The scale bar indicates amino-acid substitutions. Node support values correspond to likelihood based methods aLRT SH-like (in black) and aBayes (in magenta); support of monophyletic Wnt subfamily nodes are indicated with greater fonts size. Species abbreviations: Chordate species: Tunicates: *Botryllus schlosseri* (Bsc), *Ciona savignyi* (Csa), *Ciona robusta* (Cro; formerly *Ciona intestinallis*), *Halocynthia roretzi* (Hro), *Halocynthia aurantium* (Hau), *Mogula occulta* (Moccu), *Mogula oculata* (Mocul), *Mogula occidentalis* (Mocci), *Phallusia fumigata* (Pfu) *Phallusia mammillata* (Pma), and *Oikopleura dioica* (Odi); Cephalochordates: *Branchiostoma belcheri* (Bbe), *Branchiostoma floridae* (Bfl), *Branchiostoma lanceolatum* (Bla); Vertebrates: *Danio rerio* (Dre), *Homo sapiens* (Hsa). Non-chordates species: hemichordate *Saccoglossus kowalevskii* (Sko), annelid *Capitella teleta* (Cte), mollusk *Lottia gigantea* (Lgi), and cnidarian *Nematostella vectensis* (Nve).

gene), Fdz3/6 (3 genes) and Fdz5/8 (1 gene) subfamilies– and had only 1 sFRP (*Odi_sFRP2*), in contrast to the 5 and 3 sFRP subfamilies found in vertebrates and ascidians, respectively.

Wnts Are Expressed in All Germ-Layers During *O. dioica* Development

To explore the functional consequences of the *Wnt* losses and duplications in *O. dioica* developmental programs, we investigated the expression patterns of all *O. dioica Wnt* genes during embryogenesis and larval development by whole-mount *in situ* hybridizations (WISH). Results revealed that 5 out the identified *Wnt* genes showed complex tissue-specific expression patterns that changed throughout different developmental stages (**Figure 2**), whereas no signal was detected for two *Wnt11* paralogs, namely *Odi_Wnt11b* and *Odi_Wnt11c* (**Supplementary Figure 3**), suggesting that they may function at adult stages, or they were integrated after retrotranscription in a genomic region in which no transcription is promoted.

Because many *Wnt* genes are maternally expressed and play a role in establishing the primary axis in several metazoan species, we paid special attention to the *Wnt* expression in *O. dioica*

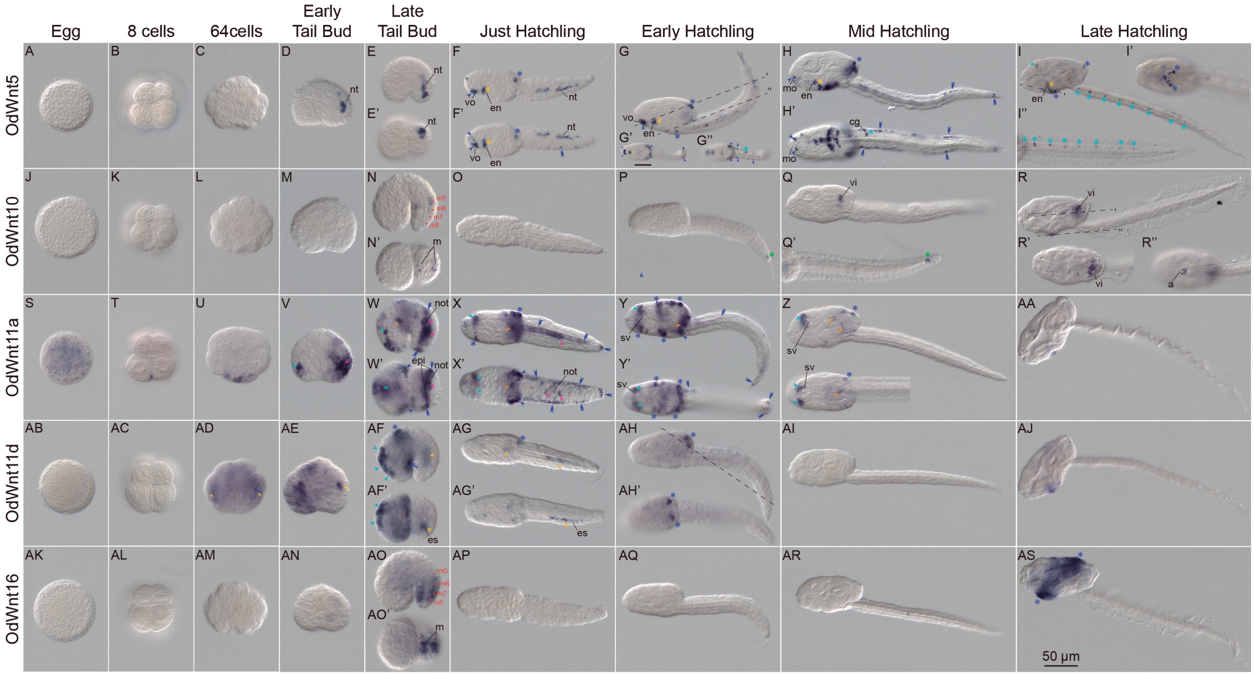


FIGURE 2 | Expression patterns of *Oikopleura dioica Wnt* genes during development. Whole-mount *in situ* hybridization in *O. dioica* eggs **(A,J,S,AB,AK)**, 8 cell embryos **(B,K,T,AC,AL)**, 64 cell embryos **(C,L,U,AD,AM)**, early tail bud embryos **(D,M,V,AE,AN)**, late tail bud embryos **(E,N,W,A,F,AO)**, just hatchlings **(F,O,X,AG,AP)**, early hatchlings **(G,P,Y,AH,AQ)**, mid hatchlings **(H,Q,Z,AI,AR)** and late hatchlings **(I,R,AA,AJ,AS)**. Upper image of each panel corresponds to left lateral view oriented anterior toward the left and dorsal toward the top. Bottom images (') are ventral views of optical cross sections at the levels of dashed lines. Red dots label muscle cell nuclei; magenta arrowheads label notochord; orange arrowheads label mesoendoderm; yellow arrowheads label endodermal strand; yellow asterisks label endostyle; light blue arrowheads label neuroectoderm and anterior brain; blue light asterisks label neural tube; green asterisk labels a cell posterior to the tip of the notochord that according to tubuline expression could be of neural condition. Blue dark arrowheads label sensory epidermis; blue dark asterisks label oikoplastic epithelium; blue dark double arrowheads label caudal epithelium and blue dark arrow labels posterior trunk epidermis. a, anus; cg, caudal ganglion; en, endostyle; epi, epidermis; es, endodermal strand, m, muscle cell pair; mo, mouth; not, notochord; nt, neural tube; sv, sensory vesicle; vi, vertical intestine; vo, ventral organ. Scale bar = 50.

eggs. WISH results revealed that from all *Wnt* genes only *Odi_Wnt11a* was part of the maternal component (**Figure 2S**). This *Wnt* paralog was, therefore, the best candidate to participate in establishing primary embryonic axis in *O. dioica* (see section following for details).

All the other *Wnt* appeared to only have zygotic expression. Of these other *Wnt*, *Odi_Wnt11d* signal was the first detected at approximately 64-cell stage. *Odi_Wnt11d* signal was detected in two internal domains, which could be endodermal derivates precursors, such as endostyle and endodermal strand (**Figure 2AD** yellow dots), while *Odi_Wnt11a* began its zygotic expression in several blastomeres in the vegetal part of the embryo (**Figure 2U**). The onset of *Odi_Wnt5* and *Odi_Wnt10* and *Odi_Wnt16* expressions occurred later, at tailbud stages, when new expression domains appeared in a highly dynamic fashion encompassing tissues derived from all three germ layers.

In the mesoderm, *Wnt* expression signal was observed in a limited number of cells of the musculature and notochord. Thus, *Odi_Wnt10* appeared to be expressed in the 5th muscle cell pair of late tailbuds (m5 in **Figure 2N**), while *Odi_Wnt16* signal appeared in the 6, 7, and 8th muscle cell pairs (m6-8 in **Figure 2AO**). At this stage, *Odi_Wnt11a* signal appeared to be restricted to the posterior third of the notochord (**Figures 2V–X** magenta arrowheads) and to the meso/endoderm in the posterior part of the trunk, anteriorly to the first cell of the notochord (**Figure 2W** orange arrowheads). This meso/endodermal expression was maintained until mid-hatchling stages, when traces of *Odi_Wnt11a* signal could be still detected in the trunk (**Figures 2X–Z** orange arrowheads).

In the endoderm, we also detected *Wnt* signal domains. From early tailbud to just-hatchling stages, *Wnt11d* labeled a group of posterior cells of the tail at the right side of notochord that correspond to the region populated by the endodermal strand (**Figures 2AE–AF** yellow arrowheads and **Supplementary Figures 4C,D**). Later, *Odi_Wnt11d* signal was detected in a region of the posterior part of the trunk where presumably the endodermal strand cells migrate (**Figure 2AG** yellow arrowheads and **Supplementary Figure 4E**). After hatch, the endostyle primordium showed a temporal expression of *Odi_Wnt11d* and *Odi_Wnt5* that was maintained up to late hatchling stages restricted to the most ventral and posterior part of the organ (**Figures 2F–I** yellow asterisk and **Supplementary Figure 4E**). At mid-hatchling stage, part of the stomach primordium was labeled by *Odi_Wnt10*, an expression domain that was maintained up to late-hatchling stages mostly restricted to the vertical intestine, to the most external part of the rectum in the opening of the anus, and to the connection between the esophagus and left stomach (**Figures 2Q,R** yellow arrowheads and **Supplementary Figure 4A**). At mid-hatchling stages, *Odi_Wnt11a* showed a faint expression in the connection between the two stomach lobes (**Figure 2Z** yellow arrowhead).

We also observed *Wnt* expression signal in derivatives of the ectoderm, including the nervous system and epidermis. At tailbud stage, *Odi_Wnt11a* labeled an anterior region of the trunk, in the presumptive area where the pharynx and the anterior brain will develop (**Figure 2V** light blue arrowheads). This expression persisted until the mid-hatchling stage, in which signal appeared labeling two bilateral domains, one adjacent to the sensory vesicle and the other at the outer part of the brain near to the epidermis, plausibly corresponding to the dorsal nerve secretomotor neurons (Olsson et al., 1990; **Figures 2W–Z** light blue arrowheads). Also, at tailbud stages, *Odi_Wnt11d* signal was observed in the neuroectoderm dorsally located to the brain, which faded in posterior stages (**Figure 2AF** light blue arrowheads), while *Odi_Wnt5* signal was observed in the posterior part of the tail, likely corresponding to the developing neural tube (**Figures 2D,E**). By hatchling stages, *Odi_Wnt5* signal exhibited a scattered pattern throughout the neural tube including the caudal ganglion (**Figures 2F–H** light blue asterisks), which became evenly distributed in late hatchlings reflecting the distribution of neurons throughout the nerve tube. *Odi_Wnt5* signal was also observed in the anterior brain of late hatchlings (**Figure 2I** light blue asterisks and arrowheads, respectively). Besides, *Odi_Wnt10* signal was also detected in nervous system, specifically in the caudal ganglion by late tailbud stage (**Supplementary Figure 4A** light blue asterisk). From early hatchling to mid hatchling stages, *Odi_Wnt10* also labeled a single cell posterior to the end of the notochord (**Figures 2P,Q** green asterisk). This position of this conspicuous cell could correspond with an α-*tubulin* positive cell described during the characterization of the nervous system in *O. diocia* (Søviknes et al., 2007), although the ontogenic nature of this cell needs further investigation (Onuma et al., 2020).

Regarding the epidermis, *Wnt* signal was observed in four different types of epidermal domains. (i) Domains in the trunk epidermis related to sensory or placode-related cells connected to the nervous system. These domains included the ventral organ and the lateral of the mouth for *Odi_Wnt5* signal (**Figures 2F–H** dark blue arrowheads), and the paired Langerhans receptor primordia in the posterior part of the trunk for *Odi_Wnt11a* signal (**Figures 2W–Y** dark blue arrowheads). (ii) Domains in the oikoplastic epithelium. The mid-dorsal domain of the oikoplast was labeled by *Odi_Wnt11a* (**Figures 2X–Z**), which also labeled the posterior-dorsal domain together with *Odi_Wnt5, Odi_Wnt11d*, and *Odi_Wnt16* (**Figures 2F–I**, **AF–AH, AS** dark blue asterisks) and the ventral oikoplast, along with *Odi_Wnt5* and *Odi_Wnt16* (**Figures 2G-I,X-Y,AS** and **Supplementary Figure 4B** dark blue asterisks). (iii) A domain in the lateral of the tail. A cell band in the equator of the tail together with the most distal epidermal cell of the tail was labeled by *Odi_Wnt5* and *Odi_Wnt11a* (**Figures 2F–H,W–Y**, dark blue double arrowheads). And iv) a domain in the posterior-ventral epidermis of the trunk, close to the separation between the trunk and the tail, was labeled by *Odi_Wnt11d* (**Figure 2AF** dark blue arrow).

Maternal Odi_Wnt11a Transcripts Are Asymmetrically Localized in the Posterior-Vegetal Region of Cleaving Embryos

As mentioned above, *Odi_Wnt11a* appeared as the only *Wnt* gene maternally transcribed in *O. dioica* (**Figure 2S**). Detail analyses during the first five cleavages of *O. dioica* embryos revealed that *Odi_Wnt11a* transcripts became asymmetrically localized in the two most posterior-vegetal cells (**Figure 3**). In eggs, both before and after fertilization, *Odi_Wnt11a* signal was uniformly distributed throughout the cytoplasm (**Figures 3A,B**). After the first division, however, *Odi_Wnt11a* signal appeared mostly accumulated nearby the contact region of the cell membranes of the two daughter cells, in the prospective vegetal-posterior pole of the embryo (**Figure 3C**). After the second division, the signal was only detected in the two posterior cells, only visible in the area toward the presumptive vegetal pole (**Figures 3D–F**). After the third division, at 8 cell-stage, *Odi_Wnt11a* signal continued restricted to the posterior region of a pair of posterior vegetal blastomeres, named B/B according to Delsman's nomenclature (Delsman, 1910; **Figures 3G–L**), equivalent to B4.1/B4.1 blastomeres according to the ascidian nomenclature system (to facilitate comparisons, ascidian nomenclature is indicated in parentheses) (Conklin, 1905; Stach et al., 2008). After the fourth and fifth divisions, *Odi_Wnt11a* signal remained in the blastomere pairs B1/B1 (B5.2/B5.2) and B11/B11 (B6.4/B6.4), which could be recognized by being the smallest ones for each division (**Figures 3M–X**). According to the fate map of *O. dioica*, these blastomeres are internalized during gastrulation, remain cleavage-arrested until the hatchling stage, and plausibly become the primordial germ cells (Fujii et al., 2008; Stach et al., 2008; Olsen et al., 2018).

Odi_Wnt11a *Knockdown*

Among the seven *Wnt* genes of *O. dioica,* we focused on *Odi_Wnt11a* because the asymmetrical distribution of its maternal component suggested a possible role in the establishment of the embryonic primary axis. To investigate the function *Odi_Wnt11a* during development, we generated knockdown animals using a DNAi approach (Omotezako et al., 2017, 2015), which consisted in the microinjection of a double strand DNA (dsDNA) against the target gene. We PCR amplified two dsDNA fragments of 206 bp and 363 bp that extended over the first exon and the 3'UTR region of the *Odi_Wnt11a*, respectively (**Figure 4A** and **Supplementary Figure 5**). These dsDNA were co-injected with an mRNA encoding for *Lifeact*-mCherry fusion protein into the ovary of pre-spawning females (Omotezako et al., 2013;see section Materials and Methods in Supporting Information for a detail description of dsDNA technique). As expected from this experimental approach, a gradient of dsDNA and mCherry was generated in the syncytial gonad from the point of injection. In the clutch from an injected female we found, therefore, both fluorescent

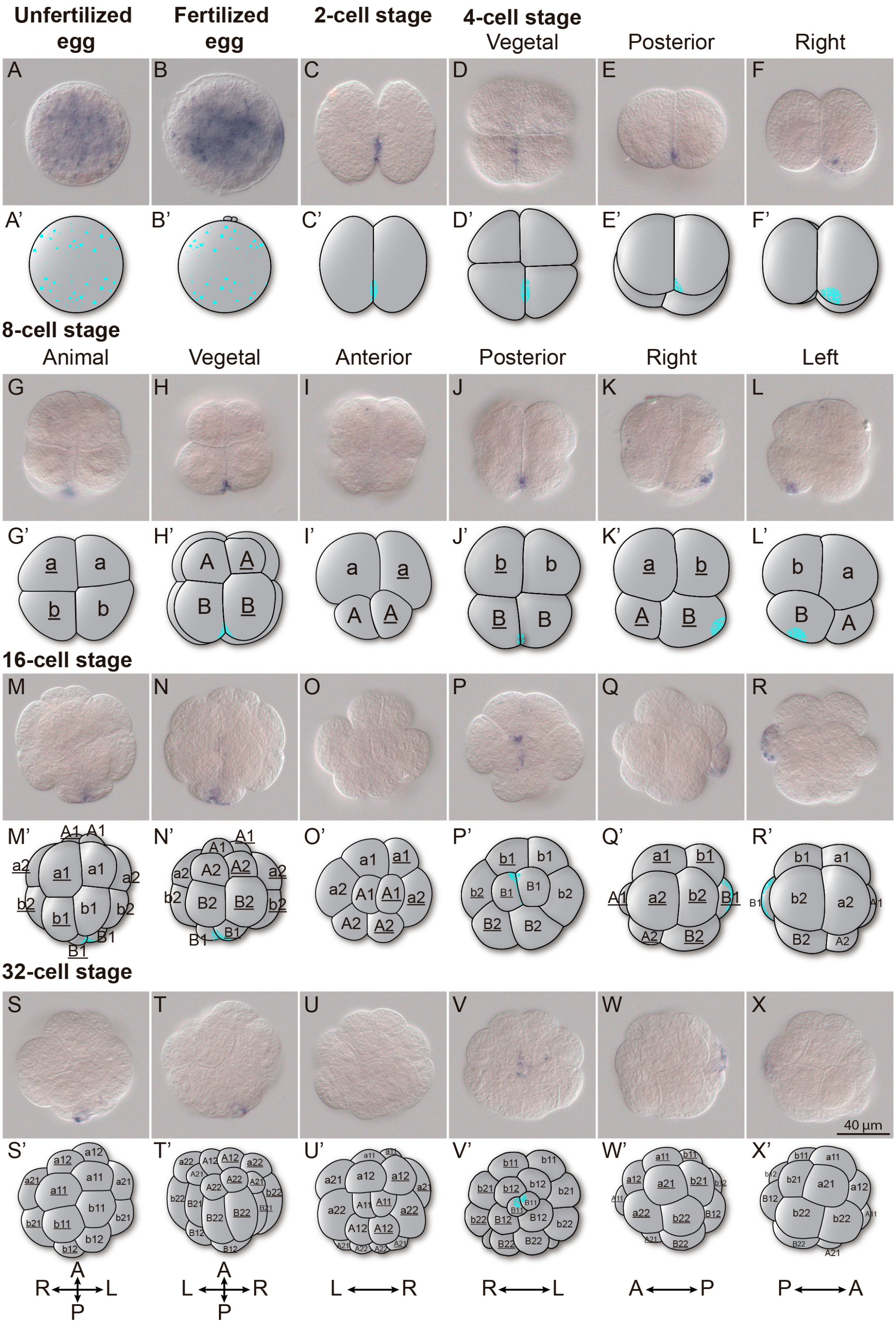

Unfertilized egg
Fertilized egg
2-cell stage
4-cell stage
Vegetal
Posterior
Right
A
B
C
D
E
F
A'
B'
C'
D'
E'
F'
8-cell stage
Animal
Vegetal
Anterior
Posterior
Right
Left
G
H
I
J
K
L
G'
H'
I'
J'
K'
L'
16-cell stage
M
N
O
P
Q
R
M'
N'
O'
P'
Q'
R'
32-cell stage
S
T
U
V
W
X
40 μm
S'
T'
U'
V'
W'
X'
R
L
A
P
L
R
A
P
L
R
R
L
A
P
P
A

FIGURE 3 | *Wnt11a* expression in cleavage stages. Whole-mount *in situ* hybridization in *O. dioica* from egg up to the 32-cell stage revealing an asymmetric localization of the *Wnt11a* maternal transcripts. **(A)** Unfertilized egg. **(B)** Fertilized egg. **(C)** 2 cells stage. **(D–F)** 4 cells stage. **(G–L)** 8 cells stage. **(M–R)** 16 cells stage. **(S–X)** 32 cells stage. Schematic representations adapted form (Fujii et al., 2008) and blastomere names according to Delsman's nomenclature are given. Embryos were viewed from various directions indicated at the top. Scale bar = 40 μm.

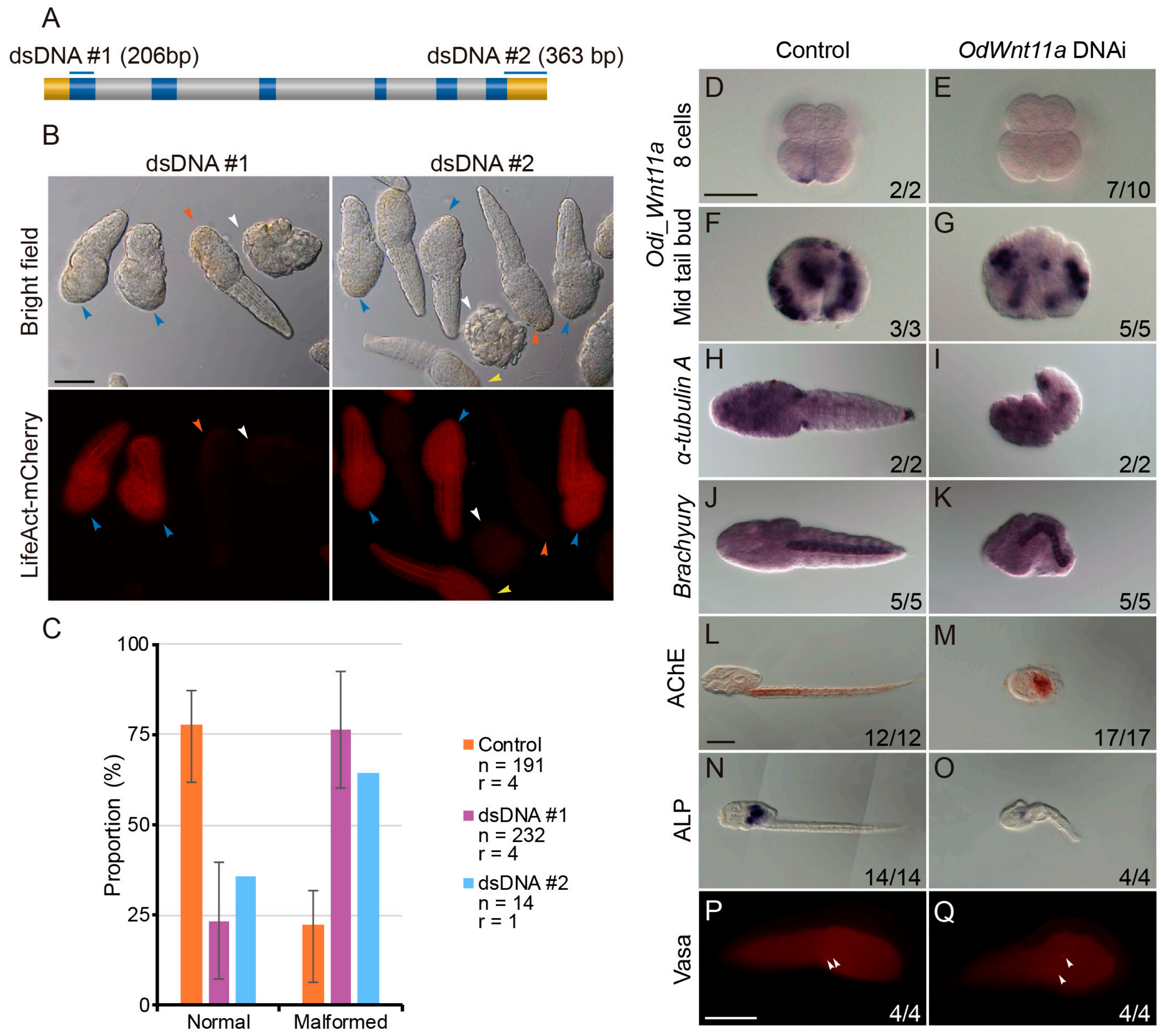


FIGURE 4 | *Wnt11a* knockdown phenotypes. **(A)** Regions targeted by each of the PCR products. Upper box is a gene model of *Wnt11a* and the regions targeted by each of the PCR products are shown by blue bars above. Blue, gray and yellow boxes indicate the ORF, intron and UTR, respectively. **(B)** Embryos co-injection with *Wnt11a* PCR products (0.2 μg/μl) and Lifeact-mCherry mRNA (0.7 μg/μl). Only *mCherry* positive embryos were examined. Phenotypes were categorized into two groups: normal and short. Two-thirds of the animals with fluorescence show shortened trunk and tail (blue arrowheads), although some animals with fluorescence did not show malformations (yellow arrowheads). Most of the animals without fluorescence developed normally (vermillion arrowheads), while some embryos without fluorescence underwent developmental arrested before hatching (white arrowhead). **(C)** Proportions of each of the phenotype of hatched larvae after injection with the three kinds of PCR products. Orange, purple, and blue bars show the result of Control (*Kaede* PCR product), DNAi #1 and DNAi #2, respectively. Number of analyzed embryos (n) and number replicas (r) are indicated in the legend. **(D–G)** Whole-mount *in situ* hybridization for *Wnt11a* in 8 cells embryos **(D,E)** and mid tail bud stage **(F,G)** injected with *Kaede* PCR product as control and *Odi_Wnt11a* DNAi #1 PCR product. **(H,I)** Whole-mount *in situ* hybridization for the neuronal marker α-*Tubulin A* in just hatchlings injected with *Kaede* PCR product and *Odi_Wnt11a* DNAi #1 PCR product. **(J,K)** Whole-mount *in situ* hybridization for the notochord marker *Brachyury* in just hatchlings injected with *Kaede* PCR product and *Odi_Wnt11a* DNAi #1 PCR product. L and M. Histochemical staining for the muscle differentiation marker acetylcholinesterase (AChE) in larvae injected with *Kaede* PCR product and *Odi_Wnt11a* DNAi #1 PCR product. **(N,O)** Histochemical staining for the endoderm differentiation marker alkaline phosphatase (ALP) in larvae injected with *Kaede* PCR product and *Odi_Wnt11a* DNAi #1 PCR product. **(P,Q)** Immunohistochemistry for the germ line marker vasa in just hatchlings injected with *Kaede* PCR product and *Odi_Wnt11a* DNAi #1 PCR product. The ratio, at the bottom of each panel, indicates the proportion of embryos with the phenotype shown. Scale bar = 50 μm.

mCherry embryos (**Figure 4B** blue and yellow arrowheads) and non-fluorescent embryos (**Figure 4B** vermillion and white arrowheads). According to previous works (Omotezako et al., 2015), we assumed that fluorescent embryos had incorporated the dsDNA, and therefore could show an altered phenotype, whereas non-fluorescent embryos should develop normally. The analysis of the phenotypes of animals from clutches of injected females with both dsDNA against *Odi_Wnt11a* showed that more than two thirds of *mCherry*-positive larvae showed similar abnormal morphologies (trunks and tails shorter than non-fluorescent siblings; **Figures 4B,C**), whereas control animals injected with a mock dsDNA (against Kaede protein) did not show such malformations. These results supported the specificity and reproducibility of the phenotype caused by both *Odi_Wnt11a* dsDNA. *In situ* hybridization of *Odi_Wnt11a* revealed a strong reduction *Wnt* signal in the posterior vegetal blastomeres (B/B) of 8-cell fluorescent embryos (**Figures 4D,E**). This reduction was, however, less strong in mid tailbud stage embryos (**Figures 4F,G**), which suggested that the dsDNA injected in the ovary was not able to completely silence zygotic transcription.

To understand the functional consequences in the *O. dioica* development of knockdown *Odi_Wnt11a* by dsDNA injection, we performed WISH and immunohistochemical analyses as well as enzymatic activity assays with several tissue-specific developmental markers, including those for neural tissue, notochord, muscle, endoderm or germ line. α-*Tubulin A* has been established as a general neuronal marker with expression in brain nerves, cerebral and caudal ganglia and nerve cord in *O. dioica* larvae (Seo et al., 2004; Søviknes et al., 2007). WISH analysis of α-*Tubulin A* showed similar expression level in the neurons of the trunk and tail when malformed (knockdown) and normal (control) larvae were compared (**Figures 4H,I**). We next analyzed the expression of *brachyury*, as a notochord marker, and the activity of acetylcholinesterase (AchE), as a muscle marker (Bassham and Postlethwait, 2000; Nishino et al., 2000; Omotezako et al., 2017), to evaluate the affectation of mesodermal tissues. Results showed that despite the malformations of the embryos, both notochord and muscle cells were present in knockdown larvae (**Figures 4J–M**). Endodermal cells were visualized in the inner region of the *O. dioica* trunk labeling the digestive tract by the activity of alkaline phosphatase (ALP) (Imai et al., 2000; Omotezako et al., 2017). ALP activity was detected in knockdown larvae, although the extension and intensity of the signal was clearly reduced (**Figures 4N,O**). Finally, because B11/B11 blastomeres, in which *Odi_Wnt11a* is asymmetrically located, are the precursor cells of the primordial germ cells (PGC), we investigated whether knocking down *Odi_Wnt11a* might affect PGC determination. We analyzed the expression of *vasa*, a germ-line specific marker (Ganot et al., 2007; Henriet et al., 2015; Olsen et al., 2018). Immunohistochemistry with an ascidian vasa anti-body (Shirae-Kurabayashi, 2006), labeled two cells in the posterior part of the trunk (**Figures 4P,Q**) in both, treated and control larvae, suggesting that maternal *Odi_Wnt11a* was not necessary for PGC determination.

In summary, the analysis of different cell- and tissue-specific markers indicated that knockdown maternal *Odi_Wnt11a* caused major morphological malformations such as shorter trunks and tails with bended notochords affecting tail elongation, as well as some impaired endodermal structures, but it did not seem to affect germ-layer specification neither overall tissue differentiation. Further investigations will be needed to establish whether maternal *Odi_Wnt11a* was not relevant for axial developmental processes or whether zygotic expression *Odi_Wnt11a*, which was only slightly affected by dsDNA injections, might be functionally compensating the maternal component.

DISCUSSION

Evolution of the Wnt Repertoire in *O. dioica*

The repertoire of *Wnt* ligands in deuterostomes –ambulacraria plus chordates– seems, in general, refractory to gene loss. In ambulacraria, for instance, the hemichordate *Saccoglossus kowalevskii* has retained the complete set of Wnt subfamilies (Darras et al., 2018), while echinoderms have all the Wnt subfamilies except the Wnt11 (Croce et al., 2006; McCauley et al., 2013; Robert et al., 2014). Similarly, two of the three chordate subphyla –cephalochordates and vertebrates– also follow this "conservative" pattern of Wnt evolution preserving all Wnt subfamilies, with the exception of WntA in vertebrates, and the duplication of several Wnt subfamilies as a product of the two rounds of whole genome duplication occurred at the base of the vertebrates (Somorjai et al., 2018). Our previous work on ascidian *Wnt* repertoire (Somorjai et al., 2018), and specially our present analysis in *O. dioica* indicates that, in contrast to all other chordates, tunicates have adopted a "liberal" pattern of evolution including several gene losses and duplications. According to our phylogenetic reconstructions, Wnt4 and Wnt8 subfamilies were lost in both ascidian and appendicularian species, and therefore, likely due to an early loss event in the last common ancestor of all tunicates. The ancestral loss of Wnt4 and Wnt8, therefore, appears as a synapomorphic trait that might have differentially influenced the evolutionary divergence of tunicates in contrast to other chordates (**Figure 5A**). In addition, the independent loss of numerous Wnt subfamilies in some groups of ascidians (i.e., Wnt1, Wnt3, and Wnt11) and in *O. dioica* (i.e., Wnt1, Wnt2, Wnt3, Wnt6, Wnt7, Wnt9, and WntA subfamilies), some of them in parallel in different tunicate lineages (i.e., Wnt1, Wnt3) suggest that the selective competence to lose Wnt genes differ between tunicates and the other chordates. *O. dioica*, which has retained members in only 4 Wnt subfamilies (i.e., Wnt5, Wnt10, Wnt11, and Wnt16; **Figure 5A**) appears as the extreme case pushing the limits of reducing the Wnt repertoire in chordates. *O. dioica,* together with the protostome gastropod *Patella vulgata* with also only 4 Wnt gene families (Prud'homme et al., 2002), would represent the minimal Wnt repertoire among all bilaterians analyzed so far.

The loss of Wnt subfamilies during the evolution of the appendicularian lineage has been accompanied by the expansion of the Wnt11 subfamily up to 4 paralogs (*Odi_Wnt11a*, *Odi_Wnt11b*, *Odi_Wnt11c*, and *Odi_Wnt11d*). The intronless structure of *Odi_Wnt11b*, *Odi_Wnt11c*, and *Odi_Wnt11d* suggest a retrotranscriptional origin of these genes from an ancestral *Odi_Wnt11a* that still retain introns (**Supplementary Figure 1**; Kaessmann et al., 2009). Often, intronless retroposed gene copies have been viewed as evolutionary dead-ends with little biological relevance due to the lack of regulatory elements. Although this may be case for *Wnt11b* and *Wnt11c* (**Supplementary Figures 3, 6**), *Wnt11d* clearly showed a specific and dynamic

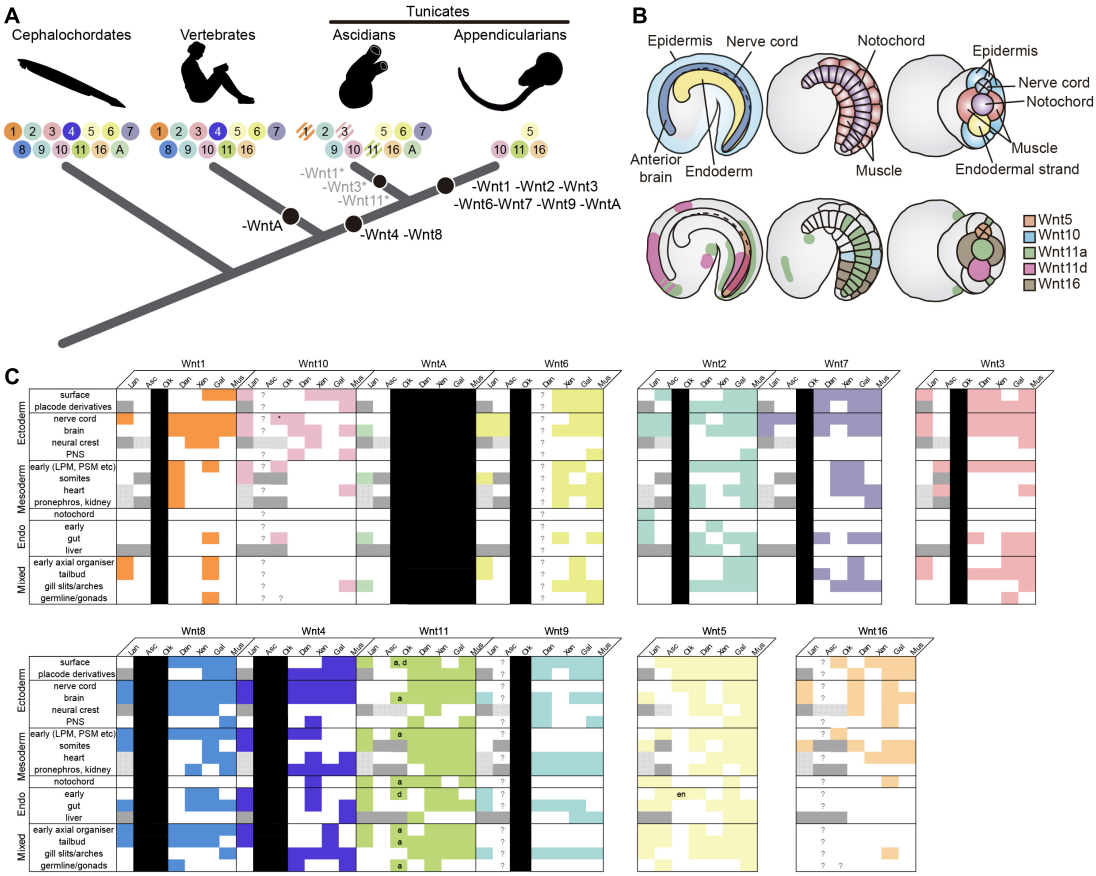

A
Tunicates
Cephalochordates
Vertebrates
Ascidians
Appendicularians
1 2 3 4 5 6 7
8 9 10 11 16 A
-Wnt1*
-Wnt3*
-Wnt11*
-WntA
-Wnt4 -Wnt8
-Wnt1 -Wnt2 -Wnt3
-Wnt6-Wnt7 -Wnt9 -WntA
B
Epidermis
Nerve cord
Notochord
Epidermis
Nerve cord
Notochord
Muscle
Anterior brain
Endoderm
Muscle
Endodermal strand
Wnt5
Wnt10
Wnt11a
Wnt11d
Wnt16
C
Wnt1
Wnt10
WntA
Wnt6
Wnt2
Wnt7
Wnt3
Wnt8
Wnt4
Wnt11
Wnt9
Wnt5
Wnt16
Lan Asc Oik Dan Xen Gal Mus
Ectoderm
surface
placode derivatives
nerve cord
brain
neural crest
PNS
Mesoderm
early (LPM, PSM etc)
somites
heart
pronephros, kidney
notochord
Endo
early
gut
liver
Mixed
early axial organiser
tailbud
gill slits/arches
germline/gonads
a, d
a
d
en
?
*

FIGURE 5 | Evolutionary comparison of the Wnt gene repertoire and expression domains between *Oikopleura dioica* and other chordates. **(A)** Evolution of the Wnt subfamily repertoire (colored circles with numbers) highlighting Wnt losses (black circles) across the chordate evolutionary tree. *O. dioica*, with only 4 Wnt gene families—*Wnt5*, *Wnt10*, *Wnt11*, and *Wnt16*–, represents the minimal *Wnt* repertoire among all chordates analyzed so far. Striped circles indicate gene losses affecting only specific groups or species (e.g., Wnt1 loss in Phlebobranchia, *Wnt11* in Molgulas, and *Wnt3* in *Botryllus*). **(B)** Schematic illustration of *Wnt* expression in *O. dioica*. Upper panel shows drawings of *O. dioica* mid tailbud embryo in lateral view (left and center) and ventral view (right), with the position of central nervous system (CNS, dark blue), epidermis (light blue), endoderm (yellow), notochord (purple), and muscles (red) depicted. Lower panels show the expression of *O. dioica*'s Wnt genes in the mid tailbud stage. **(C)** Filled colored boxes denote documented expression in Somorjai et al. (2018); white filled boxes indicate absence of, or unreported expression; question marks are reserved for genes that are present, but whose expression has never been assessed; black boxes represent lost genes; dark gray boxes denote tissues/structures that are absent in the subphylum and/or at examined stages, and stippled gray boxes highlight tissues/structures lacking clear homologs, in which therefore, expression comparison is not possible. Expression patterns for vertebrate paralogs of the same Wnt subfamily were combined. *O. dioica*'s paralogs are indicated (*a* = Odi_Wnt11a; *d* = Odi_Wnt11d). Asterisk: caudal ganglion expression (see **Supplementary Figure 4**). Lan, lancelet; Asc, ascidian; Oik, *O. dioica*; Dan, zebrafish *Danio rerio*; Xen, frog *Xenopus laevis/tropicalis*; Gal, chicken *Gallus gallus*; Mus, mouse *Mus musculus*; PNS (ectoderm), peripheral nervous system; LPM/PSM (mesoderm), lateral plate/presomitic mesoderm; en, endostyle.

expression pattern (**Figures 2AD–AH**), suggesting that it has achieved a biological role and has recruited regulatory elements that drive its expression.

Interestingly, the reduction in the repertoire of Wnt subfamilies has been accompanied by a reduction in the number of subfamilies of Wnt receptors and of antagonists (**Supplementary Figures 2B,C**), which may suggest a possible parallel gene elimination (or gene coelimination) in the Wnt activators repertoire (i.e., *Wnt* ligands and receptors) (**Supplementary Figure 2**). The conservation and expression of most of the intermediate effectors and nuclear effectors (**Supplementary Figures 2A, 6** and **Supplementary Table 1**) would indicate that the signaling pathway is totally functional in this species despite the reduction, and it will require further investigations to see how the loss of Wnt subfamilies have influenced the evolution of downstream signaling effectors.

Comparative Analysis of Wnt Expression During Embryogenesis of *O. dioica* and Other Chordates

It is generally accepted in EvoDevo that orthologous genes in the same subfamily often play conserved functions across evolutionarily distant species (Dolinski and Botstein, 2007). The comparison of the expression patterns of *O. dioica Wnt* genes with other chordate species further support this general notion (**Figures 5B,C**). *O. dioica,* however, has lost many Wnt subfamilies, and some of these losses have been possible because these subfamilies have become, somehow, dispensable. Our analysis of *O. dioica* has revealed that Wnt dispensability might be associated to synfunctionalization events (Gitelman, 2007)—i.e., one paralog acquires the expression domain of another, replacing it—leading to function shuffling when different lineages are compared (McClintock et al., 2001), or caused when a Wnt subfamily has become non-essential for *O. dioica* development. Our study of *O. dioica* Wnt genes reveals, therefore, examples of three evolutionary patterns: patterns of functional conservation, patterns of functional shuffling, and patterns of functional extinction.

Regarding the first evolutionary pattern, expression analyses suggest that many *O. dioica* Wnt orthologs conserve ancestral expression domains related to their functions in homologous structures. For instance, *Wnt11* is expressed in endodermal derivates in cephalochordates and vertebrates (Schubert et al., 2000a; Sinner et al., 2006; Cui et al., 2011; Somorjai et al., 2018), and in O. dioica Odi_*Wnt11d* is expressed in the endostyle and the endodermal strand (**Figures 2AE–AG**, **5B,C** and **Supplementary Figure 4D**), suggesting a functional conservation of Wnt11 endodermal signaling in chordates. Similarly, Odi_*Wnt11a* is expressed in the posterior part of the notochord (**Figures 2V–X**, **5B**), likewise *Wnt11* is expressed in the notochord of cephalochordates and vertebrates (Makita et al., 1998; Sasakura et al., 1998; Baranski et al., 2000; Schubert et al., 2001; Andre et al., 2015; **Figure 5C**). Besides, it has been shown that down-regulation of *Wnt11* expression produces miss elongation of the A-P axis in vertebrates (Rauch et al., 1997; Yamaguchi et al., 1999; Heisenberg et al., 2000; Tada and Smith, 2000; Nakamura et al., 2006; Niwano et al., 2009). Similarly, knocking-down *Odi_Wnt11a* produces *O. dioica* larvae with shortened trunks and tails. It can be argued, therefore, that the ancestral chordate function of *Wnt11* in endoderm and notochord has been preserved in *O. dioica,* although subfunctionalized between paralogs, that is, between endodermal *Odi_Wnt11d* and notochordal *Odi_Wnt11a*.

Regarding the second evolutionary pattern, synfunctionalization of *O. dioica Wnt* genes might appear as function shuffling events when compared with other chordate species. For example, *O. dioica* has lost Wnt6 *and* Wnt7 subfamilies (**Figure 5A**), which are expressed in the neural crest (NC) and the central nervous system (CNS) in all other chordates (Sasakura and Makabe, 2000; Schubert et al., 2001, 2000b; García-Castro et al., 2002; Imai et al., 2004; Martin et al., n.d.). *Odi_Wnt5* appears to have synfunctionalized in the *O. dioica* lineage compensating for Wnt6 and Wnt7 losses as it is expressed in the nerve cord during *O. dioica* embryogenesis (**Figures 5B,C**) (notice that *Wnt5* is not involved in the development and differentiation of the neuronal system in cephalochordates (Schubert et al., 2001), neither is expressed in the neural tube of ascidians (Miya and Nishida, 2002; Imai et al., 2004) (although the ascidian *Halocynthia roretzi* seems to transiently express *Wnt5*α in blastomeres A8.15/16 precursors of the spinal cord; Sasakura et al., 1998). Function shuffling among *Wnt5* and *Wnt6* or *Wnt7* has therefore occurred during the evolution of different chordate lineages. Function shuffling is also observed when comparing the *Wnt* genes responsible for determination of primary body axis. Whereas *Odi_Wnt11a* is the main candidate for this role in *O. dioica* (**Figure 3**), *Wnt5* in ascidians (Sasakura et al., 1998; Imai et al., 2004), *Wnt8* in zebrafish (Lu et al., 2011), and *Wnt11* in *Xenopus* (Tao et al., 2005) have been associated with this function in these species.

Regarding the third evolutionary pattern, the development of a structure that is Wnt-dependent in chordates appears to have become Wnt-independent in *O. dioica*, leading to the loss of *Wnt* genes. For example, the formation of the gill slits appears to be independent of Wnt signaling in *O. dioica* (i.e., none of the *O. dioica Wnt* genes were expressed in the gill slits), whereas Wnt signaling (together with retinoic acid signaling) appears necessary for the formation of the homologous structures in amphioxus [*Wnt9* and *WntA* are expressed in the gill arches (Onai et al., 2009; Somorjai et al., 2018) and vertebrates (*Wnt2, Wnt4, Wnt5, Wnt7, Wnt9, Wnt11,* and *Wn16* are expressed in the pharyngeal ectoderm or mesoderm of gill slits (Geetha-Loganathan et al., 2009; Choe et al., 2013; Curtin et al., n.d.; **Figure 5C**)]. It is tempting to speculate, therefore, that the loss of many Wnt subfamilies (as well as the loss of the retinoic acid signaling Martí-Solans et al., 2016) in *O. dioica* could be related with major changes in the signaling requirements of this species for the formation of particular structures, such as the development of the gill slits.

Function of the Maternal Wnt Signaling Pathway in *O. dioica*

The analysis of gene expression has shown that several Wnt signaling components of *O. dioica* (e.g., *Odi_Wnt11a*, *Odi_Fzd1/2/7 like, Odi_Fzd3/6a*, *Odi_Dvl*, *Odi_GSK3*, *Odi_*β*-catenin*. . .) are maternally expressed (Danks et al., 2013; **Supplementary Figure 6**). Among them, *Odi_Wnt11a* asymmetrically localizes in the posterior-most blastomeres during the early cleavages of development (**Figure 3**) resembling the expression of *Xenopus Wnt11* and ascidian *Wnt5* in the vegetal-posterior region (Ku and Melton, 1993; Sasakura et al., 1998).

This maternal expression of *Odi_Wnt11a* restricted to the posterior sub-cortical region of the posterior-most blastomeres suggests a role of *Wnt* signaling in A/P axis patterning (**Figure 3**). The localization of this mRNA correlates with the recently described "cytoplasm" that segregate to the posterior pole of the presumptive germ line blastomeres in *O. dioica*, called postplasm. In the *O. dioica* postplasm, maternal RNA and some morphological structures like the centrosome-attracting body (CAB) localize, reminding the postplasm of tunicate ascidians (Makabe and Nishida, 2012; Olsen et al., 2018). Interestingly, ascidian postplasm contains *Wnt5* that migrates to the posterior-most blastomeres to contribute to axis formation and cell fate determination (Prodon et al., 2007; Makabe and Nishida, 2012). Ascidian fertilized eggs injected with *Wnt5* morpholino cannot complete the gastrulation, and the asymmetric separation of the mRNAs necessary for the mesoderm endoderm segregation is also impaired (Nakamura et al., 2006; Takatori et al., 2010). When *Odi_Wnt11a* dsDNA were injected in the ovary of pre-mature *O. dioica* females, however, no defects in the cleavage and gastrulation was noticed, although a decrease in the amount of *Odi_Wnt11a* mRNA during the cleavage stages was observed (**Figure 4E**). It could be argued that the function of maternal *O. dioica* and ascidian Wnt signaling is different, or that functional inhibition by dsDNA was not strong enough to alter cleavage and gastrulation in *O. dioica* (**Figure 4**). Interestingly, it has been recently shown that knockdown of maternal β-catenin by dsDNA in *O. dioica* prevents the proper specification of the vegetal hemisphere (Omotezako et al., 2017). To fully understand the function of maternal Wnt signaling in *O. dioica* axis formation is, therefore, a challenging task that requires additional elaborate experiments such the generation of knockout lines for different Wnt genes by CRISPR that could overcome the technical limitations of lack of total penetrance of gene knockdowns.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material** or in lab web file resources: https://evodevogenomics-unibarcelona.weebly.com/lab-resources--files.html.

AUTHOR CONTRIBUTIONS

JM-S, RA, and CC designed the experiments. JM-S, HG-M, and MD-G performed genome surveys and expression analyses. JM-S was responsible of the figures and tables, and performed evolutionary inferences. JM-S performed knockdown experiments under the supervision of TO and HN. JM-S, RA, and CC were responsible for writing the manuscript. RA and CC conceived and supervise the project. All authors commented on the manuscript and agreed to its final version.

FUNDING

RA was supported by BIO2015-67358-C2-1-P, and CC was supported by BFU2016-80601-P and PID2019-110562GB-I00 from the Spanish Ministerio de Ciencia e Innovación. JM-S was supported by a research stay grant from the Montcelimar Foundation and the University of Barcelona. HN was supported by 17KT0023 and 19H03234 from Grants-in-Aid for Scientific Research (KAKENHI) from the JSPS and MEXT Japan. TO was supported by Grants-in-Aid for Scientific Research (KAKENHI) from the JSPS and MEXT Japan (24870019, 26840079, 18K06256, and 20H05946), Sumitomo Foundation (2016–2017), and the Japan Foundation for Applied Enzymology (2017, 2019). RA and CC were members of the "Grup de Recerca de la Generalitat de Catalunya" ref. 2017SGR-1665.

ACKNOWLEDGMENTS

We thank all team members of the CC and RA laboratories for fruitful discussions on Wnt signaling, gene loss, and evolution – specially Núria P. Torres-Aguila for her help in orientating the embryos in the first cleavage stages.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.700827/full#supplementary-material

REFERENCES

Albalat, R., and Cañestro, C. (2016). Evolution by gene loss. *Nat. Rev. Genet.* 17, 379–391. doi: 10.1038/nrg.2016.39

Andre, P., Song, H., Kim, W., Kispert, A., and Yang, Y. (2015). Wnt5a and Wnt11 regulate mammalian anterior-posterior axis elongation. *Development* 142, 1516–1527. doi: 10.1242/dev.119065

Baranski, M., Berdougo, E., Sandler, J. S., Darnell, D. K., and Burrus, L. W. (2000). The dynamic expression pattern of frzb-1 suggests multiple roles in chick development. *Dev. Biol.* 217, 25–41. doi: 10.1006/dbio.1999.9516

Bassham, S., and Postlethwait, J. (2000). Brachyury (T) expression in embryos of a larvacean urochordate, *Oikopleura dioica*, and the ancestral role of T. *Dev. Biol.* 220, 322–332. doi: 10.1006/dbio.2000.9647

Cañestro, C., and Postlethwait, J. H. (2007). Development of a chordate anterior-posterior axis without classical retinoic acid signaling. *Dev. Biol.* 305, 522–538. doi: 10.1016/j.ydbio.2007.02.032

Cañestro, C., Catchen, J. M., Rodríguez-Marí, A., Yokoi, H., and Postlethwait, J. H. (2009). Consequences of lineage-specific gene loss on functional evolution of surviving paralogs: ALDH1A and retinoic acid signaling in vertebrate genomes. *PLoS Genet.* 5:e1000496. doi: 10.1371/journal.pgen.1000496

Cho, S.-J., Vallès, Y., Giani, V. C., Seaver, E. C., and Weisblat, D. A. (2010). Evolutionary dynamics of the wnt gene family: a lophotrochozoan perspective. *Mol. Biol. Evol.* 27, 1645–1658. doi: 10.1093/molbev/msq052

Choe, C. P., Collazo, A., Trinh, L. A., Pan, L., Moens, C. B., and Crump, J. G. (2013). Wnt-dependent epithelial transitions drive pharyngeal pouch formation. *Dev. Cell* 24, 296–309. doi: 10.1016/j.devcel.2012.12.003

Conklin, E. G. (1905). *The Organization and Cell-Lineage of the Ascidian Egg*. Philadelphia, PA: Academy of Natural Sciences, doi: 10.5962/bhl.title.4801

Croce, J. C., Wu, S. Y., Byrum, C., Xu, R., Duloquin, L., Wikramanayake, A. H., et al. (2006). A genome-wide survey of the evolutionarily conserved Wnt pathways in the sea urchin Strongylocentrotus purpuratus. *Dev. Biol.* 300, 121–131. doi: 10.1016/j.ydbio.2006.08.045

Cui, S., Capecci, L. M., and Matthews, R. P. (2011). Disruption of planar cell polarity activity leads to developmental biliary defects. *Dev. Biol.* 351, 229–241. doi: 10.1016/j.ydbio.2010.12.041

Curtin, E., Hickey, G., Kamel, G., Davidson, A. J., and Liao, E. C. (n.d.). Zebrafish wnt9a is expressed in pharyngeal ectoderm and is required for palate and lower jaw development. *Mech. Dev* 128, 104–115. doi: 10.1016/j.mod.2010.11.003

Danks, G., Campsteijn, C., Parida, M., Butcher, S., Doddapaneni, H., Fu, B., et al. (2013). OikoBase: a genomics and developmental transcriptomics resource for the urochordate *Oikopleura dioica*. *Nucleic Acids Res.* 41, D845–D853. doi: 10.1093/nar/gks1159

Darras, S., Fritzenwanker, J. H., Uhlinger, K. R., Farrelly, E., Pani, A. M., Hurley, I. A., et al. (2018). Anteroposterior axis patterning by early canonical Wnt signaling during hemichordate development. *PLoS Biol.* 16:e2003698. doi: 10.1371/journal.pbio.2003698

Delsman, H. C. (1910). Beiträge zur entwicklungsgeschichite von *Oikopleura dioica*. *Verch. Rijksinst. Odnderz. Zee* 3, 1–24.

Dolinski, K., and Botstein, D. (2007). Orthology and functional conservation in eukaryotes. *Annu. Rev. Genet.* 41, 465–507. doi: 10.1146/annurev.genet.40.110405.090439

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Fernández, R., and Gabaldón, T. (2020). Gene gain and loss across the metazoan tree of life. *Nat. Ecol. Evol.* 4, 524–533. doi: 10.1038/s41559-019-1069-x

Ferrández-Roldán, A., Martí-Solans, J., Cañestro, C., and Albalat, R. (2019). "*Oikopleura dioica*: an emergent chordate model to study the impact of gene loss on the evolution of the mechanisms of development," in *Evo-Devo: Non-Model Species in Cell and Developmental Biology*, eds W. Tworzydlo and S. Bilinski (Cham: Springer), 63–105. doi: 10.1007/978-3-030-23459-1_4

Fujii, S., Nishio, T., and Nishida, H. (2008). Cleavage pattern, gastrulation, and neurulation in the appendicularian, *Oikopleura dioica*. *Dev. Genes Evol.* 218, 69–79. doi: 10.1007/s00427-008-0205-4

Ganot, P., Bouquet, J. M., Kallesøe, T., and Thompson, E. M. (2007). The *Oikopleura coenocyst*, a unique chordate germ cell permitting rapid, extensive modulation of oocyte production. *Dev. Biol.* 302, 591–600. doi: 10.1016/j.ydbio.2006.10.021

García-Castro, M. I., Marcelle, C., and Bronner-Fraser, M. (2002). Ectodermal Wnt function as a neural crest inducer. *Science* 297, 848–851.

Geetha-Loganathan, P., Nimmagadda, S., Antoni, L., Fu, K., Whiting, C. J., Francis-West, P., et al. (2009). Expression of WNT signalling pathway genes during chicken craniofacial development. *Dev. Dyn.* 238, 1150–1165. doi: 10.1002/dvdy.21934

Gitelman, I. (2007). Evolution of the vertebrate twist family and synfunctionalization: a mechanism for differential gene loss through merging of expression domains. *Mol. Biol. Evol.* 24, 1912–1925. doi: 10.1093/molbev/msm120

Guijarro-Clarke, C., Holland, P. W. H., and Paps, J. (2020). Widespread patterns of gene loss in the evolution of the animal kingdom. *Nat. Ecol. Evol.* 4, 519–523. doi: 10.1038/s41559-020-1129-2

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Heisenberg, C. P., Tada, M., Rauch, G. J., Saúde, L., Concha, M. L., Geisler, R., et al. (2000). Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. *Nature* 405, 76–81. doi: 10.1038/35011068

Helsen, J., Voordeckers, K., Vanderwaeren, L., Santermans, T., Tsontaki, M., Verstrepen, K. J., et al. (2020). Gene loss predictably drives evolutionary adaptation. *Mol. Biol. Evol.* 37, 2989–3002. doi: 10.1093/molbev/msaa172

Henriet, S., Sumic, S., Doufoundou-Guilengui, C., Jensen, M. F., Grandmougin, C., Fal, K., et al. (2015). Embryonic expression of endogenous retroviral RNAs in somatic tissues adjacent to the *Oikopleura germline*. *Nucleic Acids Res.* 43, 3701–3711. doi: 10.1093/nar/gkv169

Imai, K. S., Hino, K., Yagi, K., Satoh, N., and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* 131, 4047–4058. doi: 10.1242/dev.01270

Imai, K., Takada, N., Satoh, N., and Satou, Y. (2000). (beta)-catenin mediates the specification of endoderm cells in ascidian embryos. *Development* 127, 3009–3020.

Kaessmann, H., Vinckenbosch, N., and Long, M. (2009). RNA-based gene duplication: mechanistic and evolutionary insights. *Nat. Rev. Genet.* 10, 19–31. doi: 10.1038/nrg2487

Ku, M., and Melton, D. A. (1993). Xwnt-11: a maternally expressed Xenopus wnt gene. *Development* 119, 1161–1173.

Loh, K. M., van Amerongen, R., and Nusse, R. (2016). Generating cellular diversity and spatial form: Wnt signaling and the evolution of multicellular animals. *Dev. Cell* 38, 643–655. doi: 10.1016/j.devcel.2016.08.011

Lu, F.-I., Thisse, C., and Thisse, B. (2011). Identification and mechanism of regulation of the zebrafish dorsal determinant. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15876–15880. doi: 10.1073/pnas.1106801108

Makabe, K. W., and Nishida, H. (2012). Cytoplasmic localization and reorganization in ascidian eggs: role of postplasmic/PEM RNAs in axis formation and fate determination. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 501–518. doi: 10.1002/wdev.54

Makita, R., Mizuno, T., Koshida, S., Kuroiwa, A., and Takeda, H. (1998). Zebrafish wnt11: Pattern and regulation of the expression by the yolk cell and No tail activity. *Mech. Dev.* 71, 165–176. doi: 10.1016/S0925-4773(98)00013-6

Martin, A., Maher, S., Summerhurst, K., Davidson, D., and Murphy, P. (n.d.). Differential deployment of paralogous Wnt genes in the mouse and chick embryo during development. *Evol. Dev.* 14, 178–195. doi: 10.1111/j.1525-142X.2012.00534.x

Martí-Solans, J., Belyaeva, O. V., Torres-Aguila, N. P., Kedishvili, N. Y., Albalat, R., and Cañestro, C. (2016). Coelimination and survival in gene network evolution: dismantling the RA-signaling in a chordate. *Mol. Biol. Evol.* 33, 2401–2416. doi: 10.1093/molbev/msw118

Martí-Solans, J., Ferrández-Roldán, A., Godoy-Marín, H., Badia-Ramentol, J., Torres-Aguila, N. P., Rodríguez-Marí, A., et al. (2015). *Oikopleura dioica* culturing made easy: a low-cost facility for an emerging animal model in EvoDevo. *Genesis* 53, 183–193. doi: 10.1002/dvg.22800

McCauley, B. S., Akyar, E., Filliger, L., and Hinman, V. F. (2013). Expression of wnt and frizzled genes during early sea star development. *Gene Expr. Patterns* 13, 437–444. doi: 10.1016/j.gep.2013.07.007

McClintock, J. M., Carlson, R., Mann, D. M., and Prince, V. E. (2001). Consequences of Hox gene duplication in the vertebrates: an investigation of the zebrafish Hox paralogue group 1 genes. *Development* 128, 2471–2484.

Miya, T., and Nishida, H. (2002). Isolation of cDNA clones for mRNAs transcribed zygotically during cleavage in the ascidian, *Halocynthia roretzi*. *Dev. Genes Evol.* 212, 30–37.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35, W182–W185. doi: 10.1093/nar/gkm321

Nakamura, Y., Makabe, K. W., and Nishida, H. (2006). The functional analysis of Type I postplasmic/PEM mRNAs in embryos of the ascidian *Halocynthia roretzi*. *Dev. Genes Evol.* 216, 69–80. doi: 10.1007/s00427-005-0035-6

Nishino, A., Satou, Y., Morisawa, M., and Satoh, N. (2000). Muscle actin genes and muscle cells in the appendicularian, *Oikopleura longicauda*: phylogenetic relationships among muscle tissues in the urochordates. *J. Exp. Zool.* 288, 135–150.

Niwano, T., Takatori, N., Kumano, G., and Nishida, H. (2009). Wnt5 is required for notochord cell intercalation in the ascidian *Halocynthia roretzi*. *Biol. Cell* 101, 645–659. doi: 10.1042/BC20090042

Olsen, L. C., Kourtesis, I., Busengdal, H., Jensen, M. F., Hausen, H., and Chourrout, D. (2018). Evidence for a centrosome-attracting body like structure in germ-soma segregation during early development, in the urochordate *Oikopleura dioica*. *BMC Dev. Biol.* 18:4. doi: 10.1186/s12861-018-0165-5

Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* 64, 18–23.

Olsson, R., Holmberg, K., and Lilliemarck, Y. (1990). Fine structure of the brain and brain nerves of *Oikopleura dioica* (Urochordata, Appendicularia). *Zoomorphology* 110, 1–7. doi: 10.1007/BF01632806

Omotezako, T., Matsuo, M., Onuma, T. A., and Nishida, H. (2017). DNA interference-mediated screening of maternal factors in the chordate *Oikopleura dioica*. *Sci. Rep.* 7:44226. doi: 10.1038/srep44226

Omotezako, T., Nishino, A., Onuma, T. A., and Nishida, H. (2013). RNA interference in the appendicularian *Oikopleura dioica* reveals the function of the *Brachyury* gene. *Dev. Genes Evol.* 223, 261–267. doi: 10.1007/s00427-013-0438-8

Omotezako, T., Onuma, T. A., and Nishida, H. (2015). DNA interference: DNA-induced gene silencing in the appendicularian *Oikopleura dioica*. *Proc. R. Soc. Lond. B* 282:20150435. doi: 10.1098/rspb.2015.0435

Onai, T., Lin, H. C., Schubert, M., Koop, D., Osborne, P. W., Alvarez, S., et al. (2009). Retinoic acid and Wnt/beta-catenin have complementary roles in anterior/posterior patterning embryos of the basal chordate amphioxus. *Dev. Biol.* 332, 223–233. doi: 10.1016/j.ydbio.2009.05.571

Onuma, T. A., Hayashi, M., Gyoja, F., Kishi, K., Wang, K., and Nishida, H. (2020). A chordate species lacking Nodal utilizes calcium oscillation and Bmp for left–right patterning. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4188–4198. doi: 10.1073/pnas.1916858117

Onuma, T. A., Matsuo, M., and Nishida, H. (2017). Modified whole-mount in situ hybridisation and immunohistochemistry protocols without removal of the vitelline membrane in the appendicularian *Oikopleura dioica*. *Dev. Genes Evol.* 227, 367–374. doi: 10.1007/s00427-017-0588-1

Prodon, F., Yamada, L., Shirae-Kurabayashi, M., Nakamura, Y., and Sasakura, Y. (2007). Postplasmic/PEM RNAs: a class of localized maternal mRNAs with multiple roles in cell polarity and development in ascidian embryos. *Dev. Dyn.* 236, 1698–1715. doi: 10.1002/dvdy.21109

Prud'homme, B., Lartillot, N., Balavoine, G., Adoutte, A., and Vervoort, M. (2002). Phylogenetic analysis of the *Wnt* gene family: insights from lophotrochozoan members. *Curr. Biol.* 12, 1395–1400. doi: 10.1016/S0960-9822(02)01068-0

Rauch, G. J., Hammerschmidt, M., Blader, P., Schauerte, H. E., Strähle, U., Ingham, P. W., et al. (1997). Wnt5 is required for tail formation in the zebrafish embryo. *Cold Spring Harb. Symp. Quant. Biol.* 62, 227–234.

Ring, L., Neth, P., Weber, C., Steffens, S., and Faussner, A. (2014). β-Catenin-dependent pathway activation by both promiscuous "canonical" WNT3a-, and specific "noncanonical" WNT4- and WNT5a-FZD receptor combinations with strong differences in LRP5 and LRP6 dependency. *Cell. Signal.* 26, 260–267. doi: 10.1016/j.cellsig.2013.11.021

Robert, N., Lhomond, G., Schubert, M., and Croce, J. C. (2014). A comprehensive survey of wnt and frizzled expression in the sea urchin *Paracentrotus lividus*. *Genesis* 52, 235–250. doi: 10.1002/dvg.22754

Sasakura, Y., and Makabe, K. W. (2000). Ascidian Wnt-7 gene is expressed exclusively in the tail neural tube of tailbud embryos. *Dev. Genes Evol.* 210, 641–643. doi: 10.1007/s004270000108

Sasakura, Y., Ogasawara, M., and Makabe, K. W. (1998). HrWnt-5: a maternally expressed ascidian Wnt gene with posterior localization in early embryos. *Int. J. Dev. Biol. May* 42, 573–579.

Schubert, M., Holland, L. Z., and Holland, N. D. (2000a). Characterization of an amphioxus wnt gene, AmphiWnt11, with possible roles in myogenesis and tail outgrowth. *Genesis* 27, 1–5.

Schubert, M., Holland, L. Z., and Holland, N. D. (2000b). Characterization of two amphioxus Wnt genes (AmphiWnt4 and AmphiWnt7b) with early expression in the developing central nervous system. *Dev. Dyn.* 217, 205–215.

Schubert, M., Holland, L. Z., Stokes, M. D., and Holland, N. D. (2001). Three amphioxus Wnt genes (AmphiWnt3, AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Dev. Biol.* 240, 262–273.

Seo, H.-C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., Hansen, A., et al. (2004). Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* 431, 67–71. doi: 10.1038/nature02709

Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E., and Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nat. Commun.* 9:1215. doi: 10.1038/s41467-018-03667-1

Shirae-Kurabayashi, M. (2006). Dynamic redistribution of vasa homolog and exclusion of somatic cell determinants during germ cell specification in *Ciona intestinalis*. *Development* 133, 2683–2693. doi: 10.1242/dev.02446

Sinner, D., Kirilenko, P., Rankin, S., Wei, E., Howard, L., Kofron, M., et al. (2006). Global analysis of the transcriptional network controlling Xenopus endoderm formation. *Development* 133, 1955–1966. doi: 10.1242/dev.02358

Somorjai, I., Martí-Solans, J., Diaz-Gracia, M., Nishida, H., Imai, K. S., Escrivà, H., et al. (2018). Wnt evolution and function shuffling in liberal and conservative chordate genomes. *Genome Biol.* 19:98. doi: 10.1186/s13059-018-1468-3

Søviknes, A. M., Chourrout, D., and Glover, J. C. (2007). Development of the caudal nerve cord, motoneurons, and muscle innervation in the appendicularian urochordate *Oikopleura dioica*. *J. Comp. Neurol.* 503, 224–243. doi: 10.1002/cne.21376

Stach, T., Winter, J., Bouquet, J.-M., Chourrout, D., and Schnabel, R. (2008). Embryology of a planktonic tunicate reveals traces of sessility. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7229–7234. doi: 10.1073/pnas.0710196105

Tada, M., and Smith, J. C. (2000). Xwnt11 is a target of Xenopus Brachyury: regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. *Development* 127, 2227–2238.

Takatori, N., Kumano, G., Saiga, H., and Nishida, H. (2010). Segregation of germ layer fates by nuclear migration-dependent localization of Not mRNA. *Dev. Cell* 19, 589–598. doi: 10.1016/j.devcel.2010.09.003

Tao, Q., Yokota, C., Puck, H., Kofron, M., Birsoy, B., Yan, D., et al. (2005). Maternal Wnt11 activates the canonical Wnt signaling pathway required for axis formation in Xenopus embryos. *Cell* 120, 857–871. doi: 10.1016/j.cell.2005.01.013

Xu, Y.-C., and Guo, Y.-L. (2020). Less is more, natural loss-of-function mutation is a strategy for adaptation. *Plant Commun.* 1:100103. doi: 10.1016/j.xplc.2020.100103

Yamaguchi, T. P., Bradley, A., McMahon, A. P., and Jones, S. (1999). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. *Development* 126, 1211–1223.

Conflict of Interest: 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 © 2021 Martí-Solans, Godoy-Marín, Diaz-Gracia, Onuma, Nishida, Albalat and Cañestro. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 21 June 2021
doi: 10.3389/fcell.2021.704367

Evolution of Developmental Programs for the Midline Structures in Chordates: Insights From Gene Regulation in the Floor Plate and Hypochord Homologues of *Ciona* Embryos

Kouhei Oonuma[1,2], Maho Yamamoto[1], Naho Moritsugu[1], Nanako Okawa[1,2], Megumi Mukai[1,2], Miku Sotani[1,2], Shuto Tsunemi[1,2], Haruka Sugimoto[1], Eri Nakagome[1], Yuichi Hasegawa[1,2], Kotaro Shimai[1,2], Takeo Horie[3] and Takehiro G. Kusakabe[1,2]*

[1] Department of Biology, Faculty of Science and Engineering, Konan University, Kobe, Japan, [2] Institute for Integrative Neurobiology, Graduate School of Natural Science, Konan University, Kobe, Japan, [3] Shimoda Marine Research Center, University of Tsukuba, Shimoda, Japan

In vertebrate embryos, dorsal midline tissues, including the notochord, the prechordal plate, and the floor plate, play important roles in patterning of the central nervous system, somites, and endodermal tissues by producing extracellular signaling molecules, such as Sonic hedgehog (Shh). In *Ciona*, *hedgehog.b*, one of the two *hedgehog* genes, is expressed in the floor plate of the embryonic neural tube, while none of the *hedgehog* genes are expressed in the notochord. We have identified a *cis*-regulatory region of *hedgehog.b* that was sufficient to drive a reporter gene expression in the floor plate. The *hedgehog.b cis*-regulatory region also drove ectopic expression of the reporter gene in the endodermal strand, suggesting that the floor plate and the endodermal strand share a part of their gene regulatory programs. The endodermal strand occupies the same topographic position of the embryo as does the vertebrate hypochord, which consists of a row of single cells lined up immediately ventral to the notochord. The hypochord shares expression of several genes with the floor plate, including *Shh* and *FoxA*, and play a role in dorsal aorta development. Whole-embryo single-cell transcriptome analysis identified a number of genes specifically expressed in both the floor plate and the endodermal strand in *Ciona* tailbud embryos. A *Ciona* FoxA ortholog FoxA.a is shown to be a candidate transcriptional activator for the midline gene battery. The present findings suggest an ancient evolutionary origin of a common developmental program for the midline structures in Olfactores.

Keywords: ascidian, Sonic hedgehog, floor plate, endodermal strand, *Ciona intestinalis* sp. A, notochord, FoxA transcription factors, hypochord

OPEN ACCESS

Edited by:
Salvatore D'Aniello,
Stazione Zoologica Anton Dohrn, Italy

Reviewed by:
Honoo Satake,
Suntory Foundation for Life Sciences, Japan
Alberto Stolfi,
Georgia Institute of Technology, United States

****Correspondence:***
Takehiro G. Kusakabe
tgk@konan-u.ac.jp

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *02 May 2021*
Accepted: *25 May 2021*
Published: *21 June 2021*

Citation:
Oonuma K, Yamamoto M, Moritsugu N, Okawa N, Mukai M, Sotani M, Tsunemi S, Sugimoto H, Nakagome E, Hasegawa Y, Shimai K, Horie T and Kusakabe TG (2021) Evolution of Developmental Programs for the Midline Structures in Chordates: Insights From Gene Regulation in the Floor Plate and Hypochord Homologues of Ciona Embryos.
Front. Cell Dev. Biol. 9:704367.
doi: 10.3389/fcell.2021.704367

INTRODUCTION

The embryonic midline tissues, notochord, and floor plate are signaling centers that pattern vertebrate embryos (Placzek and Briscoe, 2005; Stemple, 2005). The notochord acts as an axial supportive structure and induces the floor plate in the neural tube and patterns somitic mesoderm via Sonic hedgehog (Shh) secretion. The floor plate then patterns the neural tube along the dorso-ventral axis also using the Shh signal. Thus these midline structures are central elements for construction of the vertebrate body plan.

In anamniote embryos, an endodermal rod-shaped structure, hypochord, transiently appears ventral to the notochord (Franz, 1898; Reinhardt, 1904; Gibson, 1910). Development of the hypochord also depends on signals from the notochord (Cleaver and Krieg, 1998). The hypochord was once thought to be simply a supportive structure (Stöhr, 1895; Corbo et al., 1997a) but several lines of evidence suggest that it plays a role in the positioning of the dorsal aorta (Cleaver et al., 1997; Löfberg and Collazo, 1997; Cleaver and Krieg, 1998; Eriksson and Löfberg, 2000; Hogan and Bautch, 2004), and for determination of left–right axis asymmetry (Danos and Yost, 1996; Lohr et al., 1997). Thus, transient midline tissues originating from different germ layers, the floor plate (ectoderm), the notochord (mesoderm), and the hypochord (endoderm), pattern the embryonic structure in vertebrates.

The hypochord shares expression of several genes with the floor plate and the notochord, including *Shh* and *FoxA* (Yan et al., 1995; Appel et al., 1999; Dal-Pra et al., 2011; Peyrot et al., 2011). Although their originating germ layers are different, progenitor cells of these midline tissues locate close to one another in the dorsal marginal zone, such as the Spemann organizer in amphibians and the embryonic shield in zebrafish (Shih and Fraser, 1995; Melby et al., 1996; Latimer et al., 2002; Latimer and Appel, 2006; Dal-Pra et al., 2011; Peyrot et al., 2011). These commonalities suggest a tight developmental and evolutionary connection among these midline structures. The notochord is the organ that define the phylum (or superphylum) Chordata, including vertebrates, tunicates, and cephalochordates (Kowalevsky, 1866, 1867; Yasuo and Satoh, 1993; Corbo et al., 1997a,b; Satoh et al., 2014). The ventral midline of the neural tube (nerve cord) in tunicate *Ciona* embryos expresses homologues of *Shh* (*hedgehog.b*) and *FoxA* (*FoxA.a*), and has been identified as the floor plate homologue (Corbo et al., 1997a; Takatori et al., 2002; Shi et al., 2009). By contrast, the presence of a hypochord homologue remains obscure in invertebrate chordates, although it has been suggested to be homologous with the epibranchial groove of amphioxus (Klaatsch, 1898) and a similarity between the hypochord and the endodermal strand of *Ciona* embryos has been pointed out (Corbo et al., 1997a).

Here we provide new evidence that the endodermal strand shares the gene regulatory mechanism with the floor plate in *Ciona* embryos. Functional analysis of the *cis*-regulatory region of the floor plate-specific *hedgehog.b* gene revealed its latent ability to drive transcription in the endodermal strand. Whole-embryo single-cell transcriptome analysis identified a number of genes specifically expressed in both the floor plate and the endodermal strand in *Ciona* tailbud embryos. These genes and their transcriptional regulation suggest an ancient evolutionary origin of a common developmental program for the midline structures in Olfactores. Our findings also support homology between the vertebrate hypochord and the tunicate endodermal strand.

RESULTS AND DISCUSSION

Transcriptional Activation by *Cis*-Regulatory Regions of *Ciona hedgehog.b* in the Floor Plate and Hypochord Homologues

Ciona hedgehog.b is expressed in the floor plate, but not in the notochord during embryogenesis (Takatori et al., 2002; Islam et al., 2010; **Figures 1A,B**). When the 2.6-kb upstream region of *hedgehog.b* connected with a Kaede reporter (*hedgehog.b > kaede*) was introduced into *Ciona* embryos, the expression of Kaede reporter was observed in the floor plate at the mid tailbud stage (**Figure 1D**). In addition to the expression in the floor plate, "ectopic" Kaede expression was observed in the endodermal strand of some embryos (**Figure 1D**). In contrast, no Kaede expression was observed in the notochord.

Because the reporter expression was only observed in a relatively small proportion of transfected embryos and the ectopic expression in the endodermal strand was observed (**Figure 1D**), we thought that additional *cis*-regulatory sequences might be present outside of the 2.6-kb upstream region. Comparative genomics between *Ciona intestinalis* type A and *Ciona savignyi* revealed that the first intron of *hedgehog.b* contains highly conserved non-coding regions, which could be candidates for such additional *cis*-regulatory sequences (**Figure 1C**). To test this possibility, we placed the first intron sequence upstream of the 2.6-kb genomic region in the *hedgehog.b > kaede* construct and examined Kaede reporter expression in embryos transfected with this DNA construct (**Figure 1D**). As expected, higher frequency of Kaede expression in the floor plate was observed (**Figures 1D,E**). However, the reporter expression in the endodermal strand also remained (**Figures 1D,F**). The endodermal strand is a caudal midline structure that lies immediately ventral to the notochord and its homology with the vertebrate hypochord has been proposed (Corbo et al., 1997a). Thus, the *cis*-regulatory regions of *Ciona hedgehog.b* can activate transcription in the floor plate and hypochord homologues. This observation further prompted us to test an idea that the floor plate and the endodermal strand share a developmental program including the transcriptional machinery.

Single-Cell Transcriptomic Analysis Revealed a Gene Battery Shared Among the Midline Tissues

To further investigate the shared developmental program between the floor plate and the endodermal strand, we compared gene expression profiles between these tissues by whole-embryo single-cell transcriptomics at the mid tailbud stage (**Table 1** and **Figure 2**; Horie T. et al., 2018; Horie R. et al., 2018;

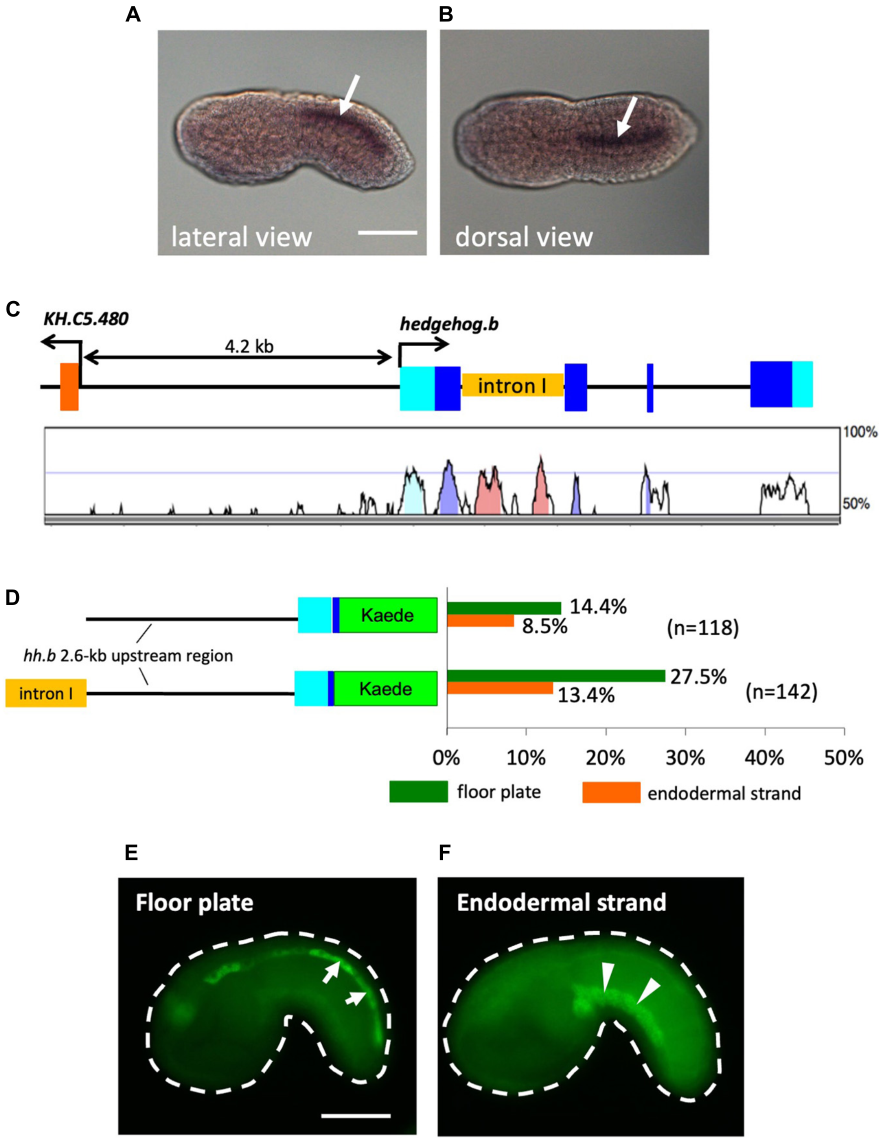


FIGURE 1 | Activity of *cis*-regulatory regions of *Ciona hedgehog.b* in the floor plate and the endodermal strand of tail bud embryos. **(A,B)** Expression pattern of *hedgehog.b* at the tailbud stage visualized by whole-mount *in situ* hybridization. Anterior is to the left. *Arrows* indicate the floor plate (ventral nerve cord). **(A)** Lateral view. **(B)** Dorsal view. **(C)** Genomic organization of the *hedgehog.b* gene. The exons are indicated by boxes in cyan (untranslated regions) and blue (coding regions). The untranscribed regions and introns are indicated by lines. A peaks-and-valleys graph below the genomic organization diagram is a VISTA profile (Frazer et al., 2004) showing the percent conservation of the nucleotide sequence of each region between *C. intestinalis* type A and *Ciona savignyi*. **(D)** The structure of the Kaede reporter constructs and the reporter expression patterns observed. The left diagrams show schematic structure of each construct. Cyan and blue boxes indicate the 5′ untranslated region and a partial coding region, respectively, of *hedgehog.b*. The orange box indicates the first intron sequence of *hedgehog.b*. Bars in the right graph show the percentage of larvae with Kaede expression in a tissue out of all electroporated embryos scored for each construct. Different colors of bars represent expression in different tissues: green, floor plate; orange, endodermal strand. Numbers in parentheses indicate the number of larvae scored for each construct. **(E,F)** Examples of mid tailbud embryos electroporated with the *hedgehog.b*(+int) > *kaede* construct showing Kaede expression in the floor plate (arrows in panel **E)** and the endodermal strand (arrowheads in panel **F)**. Scale bars, 50 μm.

Cao et al., 2019). Whole-embryo single-cell transcriptome data clearly revealed that *hedgehog.b* is expressed in the floor plate but not expressed in any other tissues, including the notochord and the endodermal strand (**Figure 2B**). Among the top 20 genes highly expressed in the endodermal strand, 8 genes were shown to be significantly enriched ($p < 0.05$) in the floor plate (**Table 1**). Of these, five genes were highly enriched ($p < 0.001$) in the floor plate (**Table 1** and **Figure 2**). These genes include *fz4* (gene model ID KH.C6.162) encoding a Frizzled4 receptor, *foxA.a* (KH.C11.313) encoding a FoxA transcription factor, KH.C2.442 encoding a solute carrier family 1 protein, KH.C5.232 encoding a tissue inhibitor of metalloproteinases 4, and KH.C4.230 encoding a calponin/transgelin family protein (*transgelin-related.b*). Interestingly, four of these genes (*fz4*,

TABLE 1 | Top 20 upregulated genes in the endodermal strand at the mid tailbud stage.

Gene Model ID	Endodermal strand		Floor plate		Similarity or predicted gene product
	$\log_2$ fold change	p-value	$\log_2$ fold change	p-value	
KH2012:KH.L41.54	6.44	3.63E-35	−1.61	1	Zinc transporter ZIP1
KH2012:KH.C4.693	5.54	1.39E-30	−2.18	1	SLIT and NTRK-like protein
KH2012:KH.C9.162	3.76	2.90E-15	−0.27	1	Regulator of G-protein signaling
KH2012:KH.C9.672	3.59	1.13E-14	0.22	1	Regulator of G-protein signaling
KH2012:KH.C1.520	3.28	1.79E-14	−0.02	1	Secreted frizzled-related protein (sFRP3/4-b)
KH2012:KH.C5.232	3.26	6.66E-13	3.41	4.04E-05	Tissue inhibitor of metalloproteinases 4
KH2012:KH.L46.15	2.79	2.65E-10	0.61	1	Uncharacterized protein
KH2012:KH.C6.162	2.88	2.69E-08	3.99	4.12E-07	Frizzled receptor (Fz4)
KH2012:KH.C4.230	2.76	4.52E-08	4.22	3.96E-09	Transgelin/Calponin/Neuronal protein 25/SM22a (tagln-r.b)
KH2012:KH.C2.378	2.81	4.75E-08	3.00	3.42E-03	brain-enriched hyaluronan-binding protein
KH2012:KH.C6.37	2.77	5.59E-08	−4.91	1	P-loop containing nucleoside triphosphate hydrolases
KH2012:KH.C9.174	3.22	7.21E-08	−3.65	1	Hypothetical protein
KH2012:KH.C3.203	2.58	9.24E-08	2.34	4.92E-02	Sulfotransferase
KH2012:KH.C2.442	2.83	2.73E-07	4.18	1.62E-07	Solute carrier family 1 (glial high affinity glutamate transporter)
KH2012:KH.C10.229	2.51	3.56E-07	−4.57	1	Phosphatidylcholine transfer protein
KH2012:KH.C2.245	2.63	7.06E-07	−0.53	1	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2
KH2012:KH.C11.313	2.53	2.85E-06	3.55	2.97E-05	Fork head/HNF-3 homologue (FoxA-a)
KH2012:KH.C9.229	2.27	4.50E-06	0.98	1	Nck-associated protein 5
KH2012:KH.L4.17	2.09	6.46E-05	−1.17	1	Zinc finger protein (Sal-like protein 1)
KH2012:KH.C3.585	2.13	7.16E-05	2.49	2.49E-02	SCRaMblase (phospholipid scramblase) family member (scrm-1)

foxA.a, KH.C5.23, and KH.C4.230) are also expressed in the notochord (**Figures 2D–G**). The expression pattern of *foxA.a* is consistent with the previously reported whole-mount *in situ* hybridization (Corbo et al., 1997a). These genes may constitute a gene battery co-regulated in the midline tissues at the mid tailbud stage.

For further analysis, we adopted KH.C4.230 as a model to investigate transcriptional regulation in the midline tissues because its expression level is relatively high and the enriched expression in the floor plate, the notochord, and the endodermal strand is strongly supported by the single-cell transcriptomic analysis (p-values, 3.96E-09, 1.49E-14, and 4.52E-08, respectively). KH.C4.230 encodes a protein belonging to the calponin/transgelin family. Calponins and transgelins are actin-associated proteins highly conserved from yeast to mammals (Prinjha et al., 1994; Goodman et al., 2003). We named KH.C4.230 as *transgelin-related.b* (*tagln-r.b*) based on the sequence similarity and genomic arrangement (**Figure 3**). Whole-mount *in situ* hybridization confirmed that *tagln-r.b* is expressed in the floor plate, the notochord, and the endodermal strand (**Figure 2H**).

In the genome of *C. intestinalis* type A, *tagln-r.b* is clustered in tandem with five other calponin/transgelin family genes within a 20-kb genomic region (**Figure 3A**). Whole-embryo single-cell transcriptome and high-throughput *in situ* hybridization data in the Ghost database (Satou et al., 2005) indicate that at least three of these *tagln-r* genes (*tagln-r.c*, *tagln-r.d*, and *tagln-r.e*) are also specifically expressed in the floor plate, the notochord, and the endodermal strand (**Supplementary Figure 1**; spatial expression patterns of *tagln-r.e* can be found at http://ghost.zool.kyoto-u.ac.jp/cgi-bin/photogetkh.cgi?inkey=CLSTR02020). Thus the clustered *tagln-r* genes are likely to be co-regulated as a member of the gene battery above mentioned.

The Role of FoxA.a as a Common Transcriptional Activator for the Midline Gene Battery

Because the expression profile of *foxA.a* (**Figure 2F**) was very similar to that of *tagln-r.b* (**Figure 2G**), FoxA.a seemed to be a good candidate for a common transcriptional activator in the midline tissues. To test this possibility, we examined distribution of the FoxA.a binding sites in the upstream of the putative transcription start sites of each of the clustered *tagln-r* genes using a set of ChIP-on-chip data of FoxA.a (Kubo et al., 2010). As expected, FoxA.a binding sites are enriched in the 5′ flanking region of each *tagln-r* gene (**Figure 3B**). To analyze the transcriptional regulatory mechanism of *tagln-r.b*, its 2.8-kb upstream region was connected with the coding sequence of Kaede (**Figure 4A**) and introduced into *Ciona* embryos. The *tagln-r.b > kaede* DNA construct recapitulated the endogenous expression pattern of *tagln-r.b*; it was expressed in the floor plate, the notochord, and the endodermal strand (**Figure 3C**), suggesting that the 2.8-kb upstream region contains *cis*-regulatory sequences sufficient for transcription in the midline tissues.

The 2.8-kb upstream region of *tagln-r.b* contains eight putative Fox binding sites (**Figure 4A**). Among these sites, three distal sites [binding sites (BS) 1–3] are conserved between *C. intestinalis* type A and *Ciona savignyi*, whereas five proximal sites (BS4-8) are not conserved. To test functional importance of putative Fox BSs, three conserved sites (BS1-3) or all eight BSs (BS1-8)

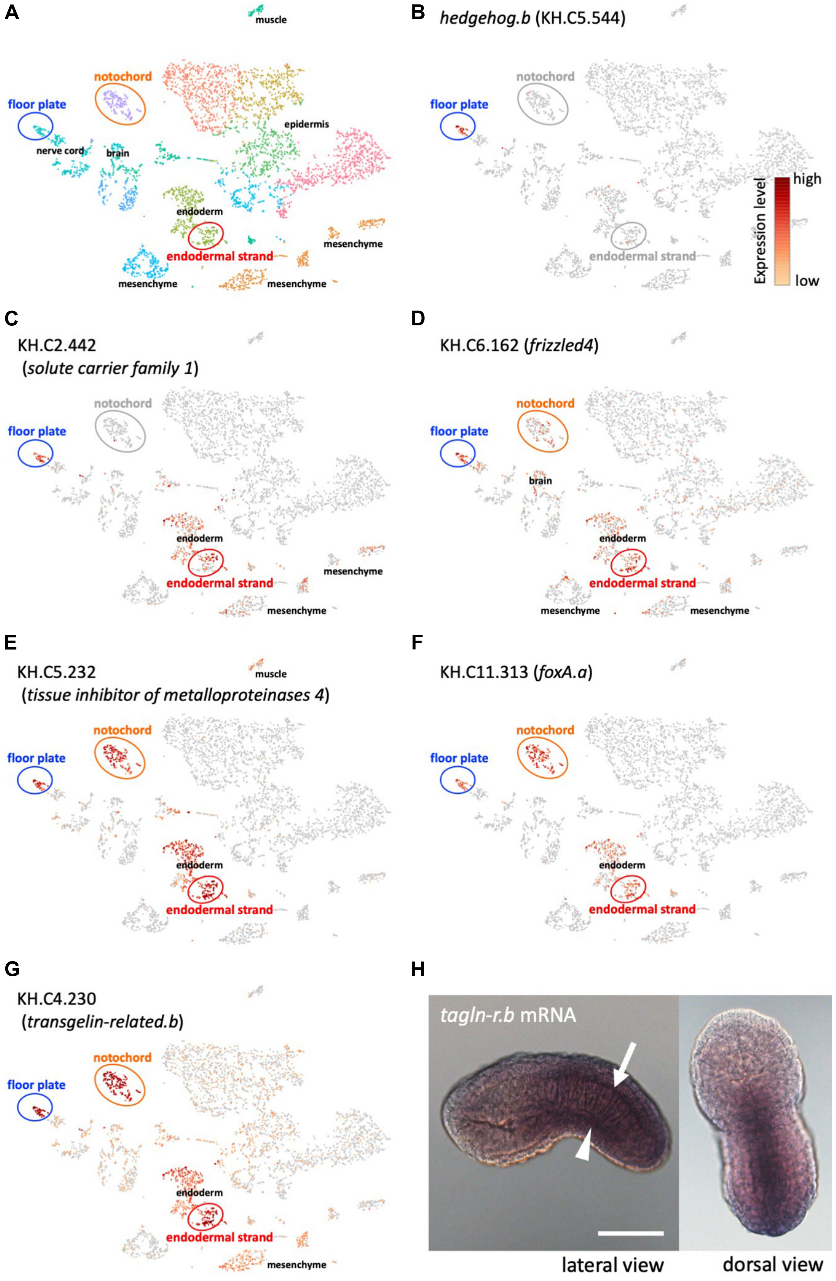


FIGURE 2 | Whole-embryo single-cell RNA-seq analysis of midline tissue-specific genes. **(A)** A t-distributed stochastic neighbor embedding (t-SNE) projection map of mid-tailbud stage embryos obtained in a previous study (Horie T. et al., 2018). Each dot corresponds to the transcriptome of a single cell, and cells possessing similar transcriptome profiles map near each other. The major tissue types in tailbud-stage embryos were identified. Identification of tissue types is based on the expression of known marker genes as previously described (Horie T. et al., 2018). Clusters of cells corresponding to the floor plate, the notochord, and the endodermal strand are encircled. **(B)** The t-SNE projection map highlighting *hedgehog.b*-expressing cells (red dots) in the floor plate. **(C–G)** The t-SNE projection map showing the expression profiles of representative genes that are specifically expressed in both the floor plate and the endodermal strand (**Table 1**). Red and orange dots indicate cells expressing KH.C2.442 **(C)**, KH.C6.162 (*frizzled4*, **D)**, KH.C5.232 **(E)**, KH.C11.313 (*foxA.a*, **F)**, and KH.C4.230 (*tagln-r.b*, **G)**. **(H)** Expression pattern of *tagln-r.b* at the tailbud stage visualized by whole-mount *in situ* hybridization. The *arrow* and the *arrowhead* indicate the floor plate and the endodermal strand, respectively. Scale bar, 50 μm.

were mutated in the *tagln-r.b* > *kaede* construct (**Figure 4A**) and the reporter expression was examined in the mid tailbud embryos transfected with these DNA constructs (**Figure 4B**). When only the conserved sites were mutated (ΔBS1-3), the reporter expression was significantly reduced in the floor plate and the endodermal strand, whereas it was moderately reduced in the notochord. Additional mutations in the non-conserved BSs (BS4-8) did not further decrease the reporter expression

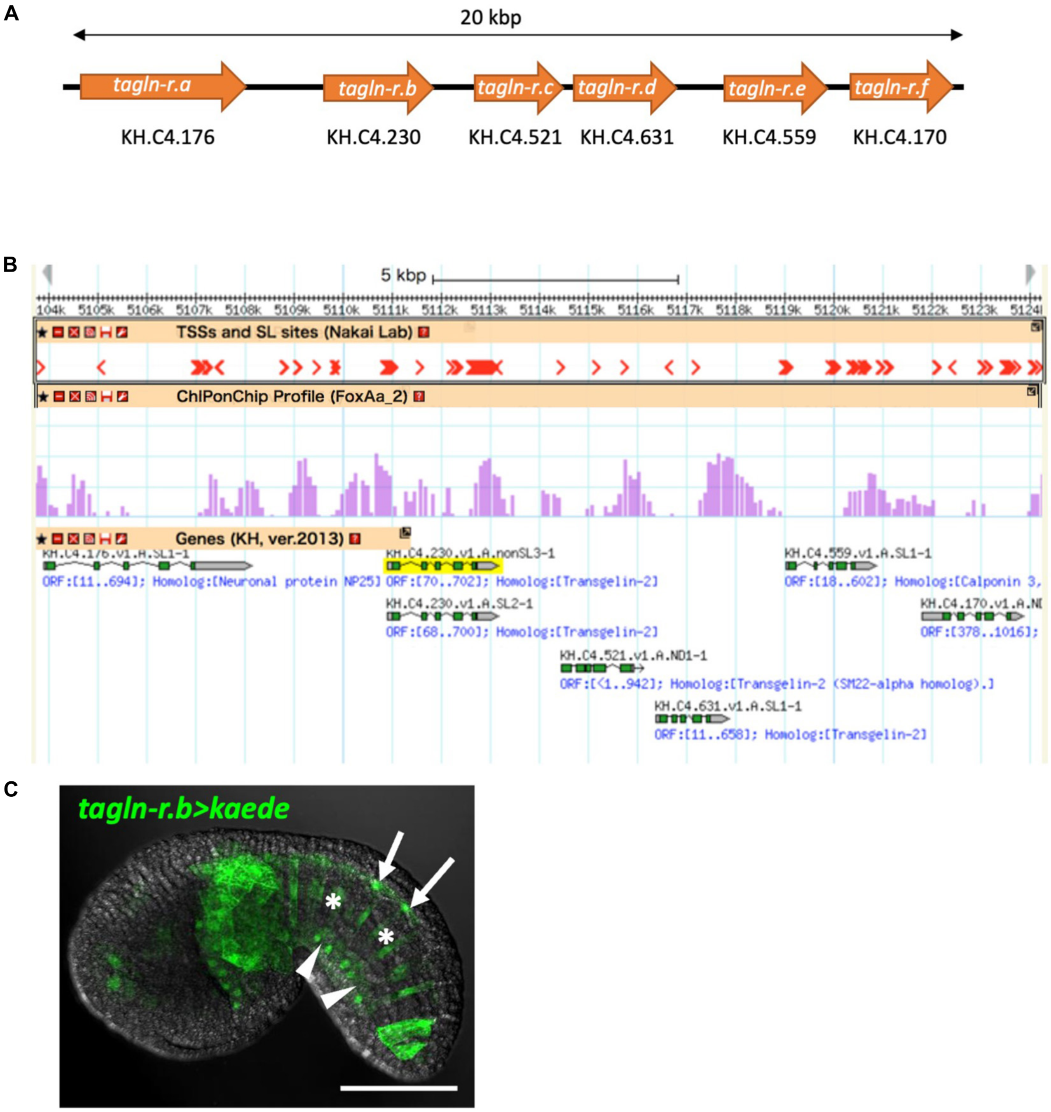

A
20 kbp
tagln-r.a
tagln-r.b
tagln-r.c
tagln-r.d
tagln-r.e
tagln-r.f
KH.C4.176
KH.C4.230
KH.C4.521
KH.C4.631
KH.C4.559
KH.C4.170
B
5 kbp
TSSs and SL sites (Nakai Lab)
ChIPonChip Profile (FoxAa_2)
Genes (KH, ver.2013)
KH.C4.176.v1.A.SL1-1
ORF:[11..694]; Homolog:[Neuronal protein NP25]
KH.C4.230.v1.A.nonSL3-1
ORF:[70..702]; Homolog:[Transgelin-2]
KH.C4.230.v1.A.SL2-1
ORF:[68..700]; Homolog:[Transgelin-2]
KH.C4.521.v1.A.ND1-1
ORF:[<1..942]; Homolog:[Transgelin-2 (SM22-alpha homolog).]
KH.C4.631.v1.A.SL1-1
ORF:[11..658]; Homolog:[Transgelin-2]
KH.C4.559.v1.A.SL1-1
ORF:[18..602]; Homolog:[Calponin 3,
KH.C4.170.v1.A.NI
ORF:[378..1016];
C
tagln-r.b>kaede

FIGURE 3 | The calponin/transgelin family gene cluster in the *Ciona* genome. **(A)** Schematic diagram of the gene cluster. Six genes, *tagln-r.a*, *tagln-r.b*, *tagln-r.c*, *tagln-r.d*, *tagln-r.e*, and *tagln-r.f*, each encoding a calponin/transgelin family protein are clustered in a 20-kb genomic region. **(B)** The transcriptional landscape of the *tagln-r* loci. Transcription start sites (TSSs) and spliced leader (SL) *trans*-splicing sites (Yokomori et al., 2016) and FoxA.a binding sites determined by ChIP-on-chip analysis (Kubo et al., 2010) were mapped on the Ghost Genome Browser (Satou et al., 2005, 2008). **(C)** Localization of Kaede reporter expressed under the control of the upstream *cis*-regulatory region of *tagln-r.b*. *Arrows*, *arrowheads*, and *asterisks* indicate the floor plate, the endodermal strand, and the notochord, respectively. Scale bar, 50 μm.

in each tissue. These results suggest that a Fox transcription factor, presumably FoxA.a, serves as a transcriptional activator of *tagln-r.b* in the midline tissues via direct interaction with the upstream region. Our observation also suggests that a greater contribution of FoxA.a to transcriptional activation of *tagln-r.b* in the floor plate and the endodermal strand than in the notochord. Because disruption of all Fox BSs in the *cis*-regulatory region of *tagln-r.b* had only slightly reduced the reporter expression in the notochord (**Figure 4**), it is plausible that Brachyury is the main activator for *tagln-r.b* in the notochord.

Among 29 Fox transcription factors identified in *C. intestinalis* type A (Imai et al., 2004; Satou et al., 2005), FoxA.a is the most plausible candidate as the transcription factor that interacts with Fox BSs in the upstream region of *tagln-r.b* for three reasons. First, as mentioned above, the ChIP-on-chip data demonstrated FoxA.a binding to the upstream region of *tagln-r.b* (Kubo et al., 2010). Second, expression patterns of *foxA.a* and *tagln-r.b* are similar to each other. Third, none of the other Fox family genes show similar expression patterns (Imai et al., 2004). In a strict sense, however, the present analysis does not exclude the possibility that a Fox transcription factor other than FoxA.a is involved in the transcriptional activation of *tagln-r.b*. To further assess the role of FoxA.a in *tagln-r.b* expression in the midline tissues, functional manipulations of FoxA.a, such as

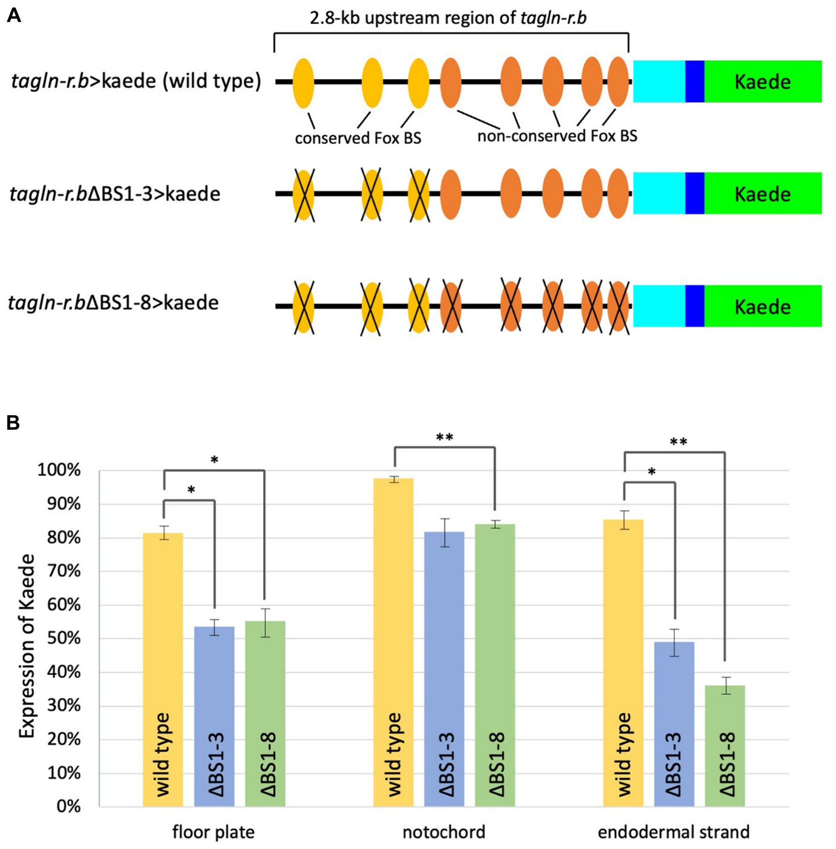


FIGURE 4 | Functional analysis of putative Fox binding sites (BSs) in the *cis*-regulatory region of *tagln-r.b*. **(A)** Schematic diagram of the Kaede reporter constructs. Cyan and blue boxes indicate the 5′ untranslated region and a partial coding region, respectively, of *tagln-r.b*. Green boxes indicate the coding sequence of Kaede. Substitution mutations were introduced into putative Fox BSs. Colored ovals indicate the putative Fox BSs and black crosses indicate mutated BSs. **(B)** Expression of the Kaede reporter in the midline tissues of tailbud embryos. Localization of Kaede was detected by immunofluorescence staining in mid tailbud (12 hpf) embryos that developed from fertilized eggs electroporated with *tagln-r.b* > *kaede* fusion constructs. Vertical bars indicate the percentage of Kaede-positive embryos. Error bars represent SEM from three independent experiments. The total number of embryos scored for each construct was 172 for *tagln-r.b* > *kaede* (wild-type), 152 for *tagln-r.b*ΔBS1-3 > *kaede*, and 147 for *tagln-r.b*ΔBS1-8 > *kaede*. Statistical analysis was carried out using the standard Student *t*-test (***P* < 0.01, **P* < 0.05).

overexpression of wild-type and a repressor form and tissue-specific knockdown, will be required in future studies.

Disruption of all Fox BSs in the *cis*-regulatory region of *tagln-r.b* did not completely abolished the reporter expression in the floor plate and the endodermal strand (**Figure 4**). This suggests that other transcription factors are involved in transactivation of *tagln-r.b*. Future identification of transcription factors that interacts with the *cis*-regulatory region of *tagln-r.b* will contribute to the elucidation of the gene regulatory networks for the development of the floor plate and the endodermal strand.

Developmental Roles of the Endodermal Strand in *Ciona* Embryos

The hypochord, transient rod-like structure situated under the notochord, is first described in embryos of elasmobranchs (Leydig, 1852). Many morphological studies on this structure were reported in embryos of lampreys, fishes, and amphibians in the late 19th and early 20th centuries (Hatta, 1893; Franz, 1898; Klaatsch, 1898; Reinhardt, 1904; Gibson, 1910). Since then, however, the hypochord has been neglected by researchers for many years, and its function remains elusive. An inductive role in the formation of the dorsal aorta has been proposed (Cleaver et al., 1997; Löfberg and Collazo, 1997; Cleaver and Krieg, 1998; Eriksson and Löfberg, 2000). Although it is uncertain whether the hypochord has a structural counterpart in embryos of higher vertebrates, a similar inductive role of the dorsal endoderm in blood vessel patterning has been proposed in avian embryos (Hogan and Bautch, 2004).

The only function of the endodermal strand known to date is its role as the precursor of the adult intestine (Hirano and Nishida, 2000; Nakazawa et al., 2013). The similarity between the hypochord and the endodermal strand prompted us to ask whether the *Ciona* endodermal strand has an inductive role similar to that of the vertebrate hypochord. In vertebrate embryos, the blood vessel precursor angioblasts migrate toward the hypochord or dorsal endoderm to form the dorsal aorta (Cleaver and Krieg, 1998; Eriksson and Löfberg, 2000; Hogan and Bautch, 2004). To test whether similar cell migration occurs in *Ciona* embryos, we labeled trunk mesenchyme cells with the photoconvertible fluorescent protein Kaede (Ando et al., 2002) and fluorescence emitted by Kaede was converted from green to red by irradiation with 405-nm violet light at 10 hpf. The *kaede* transgene was expressed using an upstream regulatory region of *Ciona pax2/5/8.a*, which could drive the reporter gene expression in trunk mesenchyme cells. These embryos were analyzed by time-laps imaging from late tailbud (12 hpf) to larval (24.5 hpf) stages (**Figure 5**). Some of the Kaede-labeled mesenchyme cells were shown to migrate into the tail along the endodermal strand (**Figure 5** and **Supplementary Video 1**). The *Ciona* endodermal strand may exert an inductive cue for the migratory mesenchyme cells, suggesting a functional similarity between the vertebrate hypochord and the *Ciona* endodermal strand.

The top 10 predominantly expressed genes in the endodermal strand include genes encoding extracellular ligands and receptors, including SLIT and NTRK-like protein (KH.C4.693), secreted frizzled-related protein (KH.C1.520), and frizzled receptor (KH.C6.162) (**Table 1**). Expression of these genes suggests an active interaction between the endodermal strand and other tissues. In zebrafish, the hypochord expresses the *frzb/sfrp3* gene that encodes a secreted frizzled-related protein (Thisse et al., 2001; Tendeng and Houart, 2006), showing a further similarity between the endodermal strand and the hypochord. Functional analysis of these genes may give insights into the role of the endodermal strand in *Ciona* embryos.

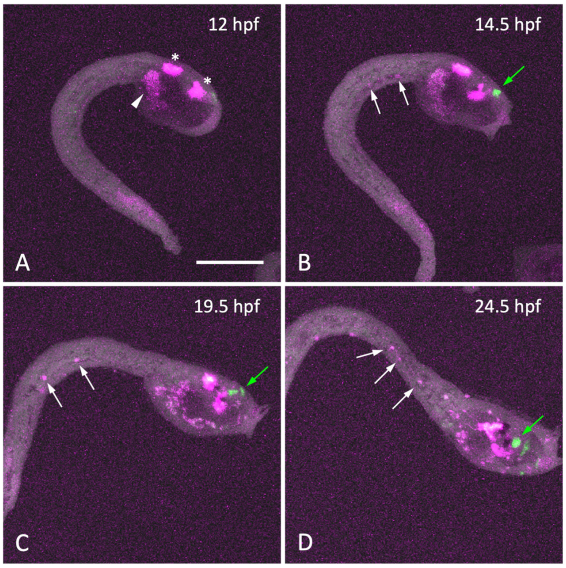


FIGURE 5 | Migration of mesenchyme cells along the endodermal strand during larval development. **(A–D)** Time-lapse fluorescent images of a late-tailbud embryo expressing Kaede under the control of the *cis*-regulatory region of *Pax2/5/8.a* at the time indicated. Kaede fluorescence was photo-converted from green to red (shown in magenta) by 405-nm laser irradiation at 12 hpf **(A)**. At 12 hpf, photo-converted Kaede fluorescence was observed in the central nervous system (*asterisks*) and mesenchyme cells (*arrowhead*) in the trunk region, whereas no cells were labeled in the tail region. As development proceeded **(B–D)**, a few cells labeled with photoconverted-Kaede appeared in the tail region and posteriorly migrated along the endodermal strand (*white arrows*). Cells synthesized Kaede after photo-conversion were labeled with green fluorescence (*green arrows* in panels **B–D)**. Scale bar, 100 μm.

Conserved Developmental Programs for Midline Tissues in Olfactores

The present findings, along with a number of previous studies, illustrate common features and the difference of midline development between vertebrates and tunicates (**Figure 6**). The gene regulatory network for notochord development in ascidian embryos has been extensively studied (Imai et al., 2006; Hotta et al., 2008; Passamaneck et al., 2009; Kubo et al., 2010; José-Edwards et al., 2015; Reeves et al., 2021). *Brachyury* is a key specifier gene for the notochord formation. FoxA.a is an upstream activator of *Brachyury* (Imai et al., 2006; Hotta et al., 2008; Kubo et al., 2010), but it also directly activates notochord-specific genes (Passamaneck et al., 2009; José-Edwards et al., 2015; Reeves et al., 2021).

Co-expression of Brachyury and FoxA family transcription factors is required for notochord development both in vertebrates and tunicates (Herrmann and Kispert, 1994; Teillet et al., 1998; Friedman and Kaestner, 2006; Imai et al., 2006; Hotta et al., 2008; Passamaneck et al., 2009; José-Edwards et al., 2015). The notochord patterns the central nervous system, somitic mesoderm, and dorsal endoderm by secreting Shh in vertebrate embryos, whereas *hedgehog* genes are not expressed in the notochord of tunicate embryos (Takatori et al., 2002). FoxA and Shh/Hedgehog.b are co-expressed in the floor plate of both vertebrates and tunicates (Tessier-Lavigne et al., 1988; Placzek et al., 1990a,b; Takatori et al., 2002; Imai et al., 2009; Dal-Pra et al., 2011; Peyrot et al., 2011). FoxA and Shh are

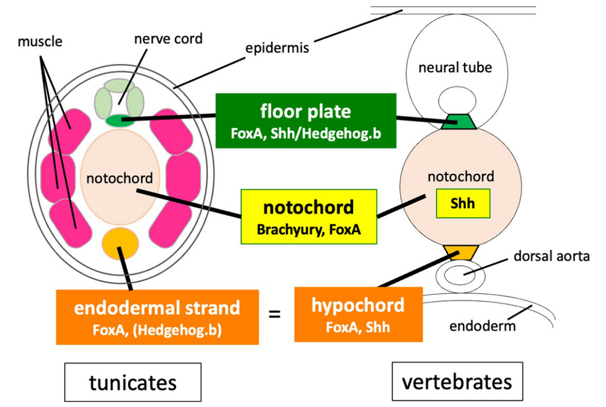


FIGURE 6 | Comparison of developmental programs for midline tissues between tunicates and vertebrates. Co-expression of Brachyury and FoxA family transcription factors is required for notochord development both in vertebrates and tunicates. Both the notochord and the floor plate secrete Shh in vertebrate embryos, whereas *hedgehog* genes are not expressed in the notochord of tunicate embryos. *FoxA* and *Shh/hedgehog.b* are co-expressed in the floor plate of both vertebrates and tunicates. *FoxA* and *Shh* are also co-expressed in the hypochord precursor. The tunicate endodermal strand expresses FoxA.a and has a latent transactivation potential of *hedgehog.b*.

also co-expressed in the hypochord or its primordium (Ruiz i Altaba, 1998; Dal-Pra et al., 2011; Peyrot et al., 2011). The tunicate endodermal strand expresses *FoxA.a* (Corbo et al., 1997a) and has a latent transactivation potential of *hedgehog.b* as shown in this study.

In conclusion, the present study suggests that the floor plate and the hypochord homologue of *Ciona* embryos share a gene battery, which is regulated by a common transcription activator FoxA.a. The FoxA transcription factor seems to be a key regulator for midline development both in ascidians and vertebrates. The endodermal strand may have an inductive role for a novel population of migratory trunk cells, which further reveals a common feature shared between the endodermal strand and the hypochord. Altogether, the present findings suggest an ancient origin of a common developmental program for and common developmental roles of the midline structures in Olfactores.

MATERIALS AND METHODS

Ciona Adults and Embryos

Mature adults of *C. intestinalis* type A (also called *Ciona robusta*) were provided by the Maizuru Fisheries Research Station of Kyoto University and by the Misaki Marine Biological Station of the University of Tokyo through the National Bio-Resource Project of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), and were maintained in indoor tanks of artificial seawater (ASW) (Marine Art BR; Tomita Pharmaceutical, Tokushima, Japan) at 18°C. The adults were also collected from the pond on the Fukae campus of Kobe University, Kobe, Japan and from the fishing harbor in Murotsu, Hyogo, Japan. Eggs and sperm were obtained surgically from the gonoducts, and the eggs were fertilized *in vitro*. After insemination, the embryos were raised in ASW containing 50 μg/ml streptomycin sulfate (S6501; Sigma-Aldrich, St. Louis, MO, United States) at 18°C.

Whole-Mount *in situ* Hybridization

The cDNA clones for *hedgehog.b* (Gene Collection ID R1CiGC41g11) and *tagln-r.b* (Gene Collection ID R1CiGC29n19) were obtained from the *Ciona* Gene Collection release 1 (Satou et al., 2002) and used as the templates to synthesize probes. To linearize the plasmid DNA for probe synthesis, cDNA clones were digested with *Xba*I (for *hedgehog.b*) or *Eco*RI (for *tagln-r.b*). Antisense RNA probes were synthesized with T7 RNA polymerase by using a DIG RNA Labeling Kit (Sigma-Aldrich, St. Louis, MI, United States). *Ciona intestinalis* type A embryos were fixed at the early tailbud stage in 4% paraformaldehyde in 0.1 M MOPS (pH 7.5) and 0.5 M NaCl at 4°C for 16 h, prior to storage in 80% ethanol at −30°C. Whole-mount *in situ* hybridization was carried out as described (Oonuma and Kusakabe, 2019).

Preparation of Reporter Constructs and Electroporation

To make the *hedgehog.b* > *kaede* plasmid, the 2.6-kb upstream region of *Ciona hedgehog.b* (Gene Model ID KH.C5.544) was amplified from the genomic DNA of *C. intestinalis* type A by PCR using a pair of nucleotide primers (5′-ATCTGCAGGGTTTGATGCACAGCAAC-3′ and 5′-ATGGATCCCCTGACCCGCATGATATGAC-3′), digested with *Pst*I and *Bam*HI, and then inserted into the *Pst*I/*Bam*HI sites of the pSP-Kaede vector (Hozumi et al., 2010). To make the *hedgehog.b*(+int) > *kaede* construct, the first intron sequence of *hedgehog.b* was amplified from the genomic DNA using a pair of nucleotide primers (5′-TTCTCGAGGCAGCAGTATGTGCCAC-3′ and 5′-CCCTGCAGCCATCCCAAGCTTCGATAAC-3′), digested with *Xho*I and *Pst*I, and then inserted into the *Xho*I/*Pst*I sites of the *hedgehog.b* > *kaede* plasmid. The *tagln-r.b* > *kaede* plasmid was made by inserting the 2.8-kb upstream region of *Ciona tagln-r.b* (Gene Model ID KH.C4.230) into the pSP-Kaede plasmid using an In-Fusion HD Cloning kit (Takara Bio, Japan). The 2.8-kb upstream region of *tagln-r.b* was amplified from the genomic DNA by PCR using a pair of nucleotide primers (5′-AAACTCGAGTCACACGAATTAAGCAAAGC-3′ and 5′-TTTTTCTCGTTGCGCCAT-3′). To generate mutant constructs, *tagln-r.b*ΔBS1-3 > *kaede* and *tagln-r.b*ΔBS1-8 > *kaede*, putative Fox binding sites (RYAAAYA; Chen et al., 2016) were mutagenized by the PCR-based method as previously described (Oonuma and Kusakabe, 2019). Oligonucleotide primers used for the mutagenesis of fox binding sites (BS1-8) were: 5′-GGTACGgcccaAAGCAGGAATTTTAATAGCAGT-3′ and 5′-CTGCTTtgggcCGTACCTTTACCTTACTGGGTGG-3′ for BS1; 5′-GCTTCTcgggtTCTTGCCAAATAAGGCGA-3′ and 5′-GCAAGAacccgAGAAGCACGAAGCAAATTC-3′ for BS2; 5′-AACTGTTtgggtTCTTGGGCGAGCTAAGC-3′ and 5′-CCCAAGAacccaAACAGTTTCATTGAAAGAGCC-3′ for BS3; 5′-GCAAAAGacccgATTCGTGCGACGGATTC-3′ and 5′-CACGAATcgggtCTTTTGCTCTCCCATGCA-3′ for BS4; 5′-CCTAGATcgggcTCGTACAACAGTTTGACGTAAGTTC-3′ and 5′-TGTACGAgcccgATCTAGGCGTATTTCCACACG-3′ for BS5; 5′-TGTTATGgcccgACTCCATTCGTTCAACTTTCTAGA-3′ and 5′-GGAGTcgggcCATAACACCATACCTGTCGCGCG-3′ for BS6; 5′-GCGTTTtgggcCGTTTGATTGATAAATGTACGTAAGAGA-3′ and 5′-CAAACGgcccaAAACGCATTTAAAAGCCAGTT-3′ for BS7; 5′-CCTCATAGacccaAGCGAATCCATTGTCAAGTC-3′ and 5′-ATTCGCTtgggtCTATGAGGAGTATAGGCGAGGTG-3′ for BS8. To make the *pax2/5/8.a* > *kaede* plasmid, the 4.4-kb upstream region of *Ciona pax2/5/8.a* (Gene Model ID KH.S545.1/KH.S1363.2; Imai et al., 2004) was amplified from the genomic DNA of *C. intestinalis* type A by PCR using a pair of nucleotide primers (5′-CGACTCTAGAGGATCCGTGATTGTTACGTGG-3′ and 5′-TGGGGATCAGCAATGGATCCCCTTGCGGCC-3′), digested with *Bam*HI, and then inserted into the *Bam*HI site of the pSP-Kaede vector. Plasmid DNA constructs were electroporated into fertilized *Ciona* eggs as described by Corbo et al. (1997b).

Immunofluorescence Staining

Immunofluorescent staining was carried out according to the method described by Nishitsuji et al. (2012). To visualize the localization of Kaede, a rabbit anti-Kaede polyclonal antibody

(PM012; Medical & Biological Laboratories, Nagoya, Japan; for Kaede) was diluted 1:1000 in 10% goat serum in T-PBS (0.1% Triton X-100 in PBS) and used as the primary antibody. The secondary antibody was an Alexa Fluor 488-conjugated anti-rabbit IgG (A11008; Thermo Fisher Scientific, Waltham, MA, United States). Fluorescent images were obtained by using a laser scanning confocal microscope (FV1200 IX83; Olympus, Tokyo, Japan). Confocal images were collected at 1-μm intervals in the *z*-axis.

Whole-Embryo Single-Cell Transcriptomic Analysis

A published single-cell transcriptome dataset of mid-tailbud embryos obtained using the 10x Genomics Chromium system (Horie T. et al., 2018; Horie R. et al., 2018; Cao et al., 2019) was used to analyze expression profiles of genes in the midline tissues. The dataset is available through GEO (GSE120035): https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120035. The t-distributed stochastic neighbor embedding (t-SNE) analysis was performed using the Loupe Cell Browser 3.1.1 software (10x Genomics, Pleasanton, CA, United States). The processed data in a Loupe Cell Browser file (.cloupe) is available through the Mendeley data repository: http://dx.doi.org/10.17632/n4pxpr28cb.1. Differentially expressed genes were identified and ranked by statistical significance as previously described (Horie T. et al., 2018).

Time-Lapse Live Imaging and Photo-Conversion of Kaede

Embryos electroporated with *pax2/5/8.a > kaede* were reared in ASW and mounted on a glass slide with ASW containing 1.5% methylcellulose at 10 hpf. Photoconversion of Kaede was performed as described (Oonuma et al., 2016). Fluorescent images were taken every 15 min for 12.5 h at 18°C by using a laser scanning confocal microscope (FV1200 IX83; Olympus, Japan). Confocal images were collected at 1-μm intervals in the *z*-axis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

TK conceived the project and wrote the manuscript. KO, KS, TH, and TK designed the experiments. KO, MY, NM, NO, MM, ST, HS, EN, YH, and KS performed the experiments. TH, ST, MS, and TK analyzed the single-cell RNA-seq data. TH provided the essential materials. KO, NO, and TK analyzed and interpreted the data. KO, TH, and TK edited the manuscript. All authors reviewed the manuscript.

FUNDING

This study was supported in part by Grants-in-Aid for Scientific Research from JSPS to TK (16H04724 and 19H03213), KO (17K15130 and 19K16150), and TH (16K07433 and 19H03204). This study was also supported in part by research grants from the Takeda Science Foundation and the Hirao Taro Foundation of KONAN GAKUEN for Academic Research.

ACKNOWLEDGMENTS

We thank the National BioResource Project of MEXT and all members of the Maizuru Fisheries Research Station and Yutaka Satou Lab of Kyoto University and the Misaki Marine Biological Station of the University of Tokyo for providing us with *C. intestinalis* adults. We also thank the Graduate School of Maritime Sciences, Kobe University for generously allowing us to collect *C. intestinalis* on the campus.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.704367/full#supplementary-material

Supplementary Video 1 | Time-lapse of an embryo showing migration of mesenchyme cells labeled with photo-converted Kaede shown in **Figure 5**.

REFERENCES

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12651–12656. doi: 10.1073/pnas.202320599

Appel, B., Fritz, A., Westerfield, M., Grunwald, D. J., Eisen, J. S., and Riley, B. B. (1999). Delta-mediated specification of midline cell fates in zebrafish embryos. *Curr. Biol.* 9, 247–256. doi: 10.1016/s0960-9822(99)80113-4

Cao, C., Lemaire, L. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., et al. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571, 349–354. doi: 10.1038/s41586-019-1385-y

Chen, X., Ji, Z., Webber, A., and Sharrocks, A. D. (2016). Genome-wide binding studies reveal DNA binding specificity mechanisms and functional interplay amongst Forkhead transcription factors. *Nucleic Acids Res.* 44, 1566–1578. doi: 10.1093/nar/gkv1120

Cleaver, O., and Krieg, P. A. (1998). VEGF mediates angioblast migration during development of the dorsal aorta in *Xenopus*. *Development* 125, 3905–3914. doi: 10.1242/dev.125.19.3905

Cleaver, O., Tonissen, K. F., Saha, M. S., and Krieg, P. A. (1997). Neovascularization of the *Xenopus* embryo. *Dev. Dyn.* 210, 66–77.

Corbo, J. C., Erives, A., Di Gregorio, A., Chang, A., and Levine, M. (1997a). Dorsovental patterning of the vertebrate neural tube is conserved in a protochordate. *Development* 124, 2335–2344. doi: 10.1242/dev.124.12.2335

Corbo, J. C., Levine, M., and Zeller, R. W. (1997b). Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, *Ciona intestinalis*. *Development* 124, 589–602. doi: 10.1242/dev.124.3.589

Dal-Pra, S., Thisse, C., and Thisse, B. (2011). FoxA transcription factors are essential for the development of dorsal axial structures. *Dev. Biol.* 350, 484–495. doi: 10.1016/j.ydbio.2010.12.018

Danos, M. C., and Yost, H. J. (1996). Role of notochord in specification of cardiac left-right orientation in zebrafish and *Xenopus*. *Dev. Biol.* 177, 96–103. doi: 10.1006/dbio.1996.0148

Eriksson, J., and Löfberg, J. (2000). Development of the hypochord and dorsal aorta in the zebrafish embryo (*Danio rerio*). *J. Morphol.* 244, 167–176. doi: 10.1002/(sici)1097-4687(200006)244:3<167::aid-jmor2>3.0.co;2-j

Franz, K. (1898). Über die entwicklung von hypochorda und *Ligamentum longitudinale ventrale* bei Teleostiern. *Morphol. Jahrb.* 25, 143–155.

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32, W273–W279.

Friedman, J. R., and Kaestner, K. H. (2006). The Foxa family of transcription factors in development and metabolism. *Cell Mol. Life Sci.* 63, 2317–2328. doi: 10.1007/s00018-006-6095-6

Gibson, W. T. (1910). The development of the hypochord in *Raia batis*; with a note upon the occurrence of the epibranchial groove in amniote embryos. *Anat. Anz.* 35, 407–428.

Goodman, A., Goode, B. L., Matsudaira, P., and Fink, G. R. (2003). The *Saccharomyces cerevisiae* calponin/transgelin homolog Scp1 functions with fimbrin to regulate stability and organization of the actin cytoskeleton. *Mol. Biol. Cell* 14, 2617–2629. doi: 10.1091/mbc.e03-01-0028

Hatta, S. (1893). On the formation of the germinal layers in Petromyzon. *J. Coll. Sci. Imp. Univ. Jpn.* 5, 129–147.

Herrmann, B. G., and Kispert, A. (1994). The *T* genes in embryogenesis. *Trends Genet.* 10, 280–286.

Hirano, T., and Nishida, H. (2000). Developmental fates of larval tissues after metamorphosis in the ascidian, *Halocynthia roretzi*. II. Origin of endodermal tissues of the juvenile. *Dev. Genes Evol.* 210, 55–63. doi: 10.1007/s004270050011

Hogan, K. A., and Bautch, V. L. (2004). Blood vessel patterning at the embryonic midline. *Curr. Top. Dev. Biol.* 62, 55–85. doi: 10.1016/s0070-2153(04)62003-5

Horie, R., Hazbun, A., Chen, K., Cao, C., Levine, M., and Horie, T. (2018). Shared evolutionary origin of vertebrate neural crest and cranial placodes. *Nature* 560, 228–232. doi: 10.1038/s41586-018-0385-7

Horie, T., Horie, R., Chen, K., Cao, C., Nakagawa, M., Kusakabe, T. G., et al. (2018). Regulatory cocktail for dopaminergic neurons in a protovertebrate identified by whole-embryo single-cell transcriptomics. *Genes Dev.* 32, 1297–1302. doi: 10.1101/gad.317669.118

Hotta, K., Takahashi, H., Satoh, N., and Gojobori, T. (2008). Brachyury-downstream gene sets in a chordate, *Ciona intestinalis*: integrating notochord specification, morphogenesis and chordate evolution. *Evol. Dev.* 10, 37–51. doi: 10.1111/j.1525-142x.2007.00212.x

Hozumi, A., Kawai, N., Yoshida, R., Ogura, Y., Ohta, N., Satake, H., et al. (2010). Efficient transposition of a single *Minos* transposon copy in the genome of the ascidian *Ciona intestinalis* with a transgenic line expressing transposase in eggs. *Dev. Dyn.* 239, 1076–1088. doi: 10.1002/dvdy.22254

Imai, K. S., Hino, K., Yagi, K., Satoh, N., and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* 131, 4047–4058. doi: 10.1242/dev.01270

Imai, K. S., Levine, M., Satoh, N., and Satou, Y. (2006). Regulatory blueprint for a chordate embryo. *Science* 312, 1183–1187. doi: 10.1126/science.1123404

Imai, K. S., Stolfi, A., Levine, M., and Satou, Y. (2009). Gene regulatory networks underlying the compartmentalization of the *Ciona* central nervous system. *Development* 136, 285–293. doi: 10.1242/dev.026419

Islam, A. F. M. T., Moly, P. K., Miyamoto, Y., and Kusakabe, T. G. (2010). Distinctive expression patterns of Hedgehog pathway genes in the *Ciona intestinalis*. *Zool. Sci.* 27, 84–90. doi: 10.2108/zsj.27.84

José-Edwards, D. S., Oda-Ishii, I., Kugler, J. E., Passamaneck, Y. J., Katikala, L., Nibu, Y., et al. (2015). Brachyury, Foxa2 and the *cis*-regulatory origins of the notochord. *PLoS Genet.* 11:e1005730. doi: 10.1371/journal.pgen.1005730

Klaatsch, H. (1898). Zur Frage nach der morphologiscen Bedeutung der Hypochorda. *Morphol. Jahrb.* 25, 156–169.

Kowalevsky, A. (1866). Entwicklungsgeschichte der einfachen Ascidien. *Mem. l'Acad. St. Petersbourg.* 7, 11–19.

Kowalevsky, A. (1867). Entwicklungsgeschichte des Amphioxus lanceolatus. *Mem. l'Acad. St. Petersbourg* 11, 1–17.

Kubo, A., Suzuki, N., Yuan, X., Nakai, K., Satoh, N., Imai, K. S., et al. (2010). Genomic *cis*-regulatory networks in the early *Ciona* intestinalis embryo. *Development* 137, 1613–1623. doi: 10.1242/dev.046789

Latimer, A. J., and Appel, B. (2006). Notch signaling regulates midline cell specification and proliferation in zebrafish. *Dev. Biol.* 298, 392–402. doi: 10.1016/j.ydbio.2006.05.039

Latimer, A. J., Dong, X., Markov, Y., and Appel, B. (2002). Delta-Notch signaling induces hypochord development in zebrafish. *Development* 129, 2555–2563. doi: 10.1242/dev.129.11.2555

Leydig, F. (1852). *Beträge zur Mikroskopischen Anatomie und Entwicklungsgeschichte der rochen und Haie*. Leipzig: Wilhelm Engelmann.

Löfberg, J., and Collazo, A. (1997). Hypochord, an enigmatic embryonic structure: study of the axolotl embryo. *J. Morphol.* 232, 57–66. doi: 10.1002/(sici)1097-4687(199704)232:1<57::aid-jmor3>3.0.co;2-l

Lohr, J. L., Danos, M. C., and Yost, H. J. (1997). Left-right asymmetry of a nodal-related gene is regulated by dorsoanterior midline structures during *Xenopus* development. *Development* 124, 1465–1472. doi: 10.1242/dev.124.8.1465

Melby, A. E., Warga, R. M., and Kimmel, C. B. (1996). Specification of cell fates at the dorsal margin of the zebrafish gastrula. *Development* 122, 2225–2237. doi: 10.1242/dev.122.7.2225

Nakazawa, K., Yamazawa, T., Moriyama, Y., Ogura, Y., Kawai, N., Sasakura, Y., et al. (2013). Formation of the digestive tract in *Ciona intestinalis* includes two distinct morphogenic processes between its anterior and posterior parts. *Dev. Dyn.* 242, 1172–1183. doi: 10.1002/dvdy.24009

Nishitsuji, K., Horie, T., Ichinose, A., Sasakura, Y., Yasuo, H., and Kusakabe, T. G. (2012). Cell lineage and *cis*-regulation for a unique GABAergic/glycinergic neuron type in the larval nerve cord of the ascidian *Ciona intestinalis*. *Dev. Growth Differ.* 54, 177–186. doi: 10.1111/j.1440-169x.2011.01319.x

Oonuma, K., and Kusakabe, T. G. (2019). Spatio-temporal regulation of *Rx* and mitotic patterns shape the eye-cup of the photoreceptor cells in *Ciona*. *Dev. Biol.* 445, 245–255. doi: 10.1016/j.ydbio.2018.11.011

Oonuma, K., Tanaka, M., Nishitsuji, K., Kato, Y., Shimai, K., and Kusakabe, T. G. (2016). Revised lineage of larval photoreceptor cells in *Ciona* reveals archetypal collaboration between neural tube and neural crest in sensory organ formation. *Dev. Biol.* 420, 178–185. doi: 10.1016/j.ydbio.2016.10.014

Passamaneck, Y. J., Katikala, L., Perrone, L., Dunn, M. P., Oda-Ishii, I., and Di Gregorio, A. (2009). Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian *Ciona intestinalis*. *Development* 136, 3679–3689. doi: 10.1242/dev.038141

Peyrot, S. M., Wallingford, J. B., and Harland, R. M. (2011). A revised model of *Xenopus* dorsal midline development: differential and separable requirements for Notch and Shh signaling. *Dev. Biol.* 352, 254–266. doi: 10.1016/j.ydbio.2011.01.021

Placzek, M., and Briscoe, J. (2005). The floor plate: multiple cells, multiple signals. *Nat. Rev. Neurosci.* 6, 230–240. doi: 10.1038/nrn1628

Placzek, M., Tessier-Lavigne, M., Jessell, T., and Dodd, J. (1990a). Orientation of commissural axons in vitro in response to a floor plate-derived chemoattractant. *Development* 110, 19–30. doi: 10.1242/dev.110.1.19

Placzek, M., Tessier-Lavigne, M., Yamada, T., Jessell, T., and Dodd, J. (1990b). Mesodermal control of neural cell identity: floor plate induction by the notochord. *Science* 250, 985–988. doi: 10.1126/science.2237443

Prinjha, R. K., Shapland, C. E., Hsuan, J. J., Totty, N. F., Mason, I. J., and Lawson, D. (1994). Cloning and sequencing of cDNAs encoding the actin cross-linking protein transgelin defines a new family of actin-associated proteins. *Cell Motil. Cytoskeleton* 28, 243–255. doi: 10.1002/cm.970280307

Reeves, W. M., Shimai, K., Winkley, K. M., and Veeman, M. T. (2021). Brachyury controls *Ciona* notochord fate as part of a feed-forward network. *Development* 148:dev195230.

Reinhardt, A. (1904). Die Hypochord bei Salamandra maculosa. *Morphol. Jahrb.* 32, 195–231.

Ruiz i Altaba, A. (1998). Combinatorial Gli gene function in floor plate and neuronal inductions by Sonic hedgehog. *Development* 125, 2203–2212. doi: 10.1242/dev.125.12.2203

Satoh, N., Rokhsar, D., and Nishikawa, T. (2014). Chordate evolution and the three-phylum system. *Proc. R. Soc. B* 281:20141729. doi: 10.1098/rspb.2014.1729

Satou, Y., Kawashima, T., Shoguchi, E., Nakayama, A., and Satoh, N. (2005). An integrated database of the ascidian, *Ciona intestinalis*: towards functional genomics. *Zool. Sci.* 22, 837–843. doi: 10.2108/zsj.22.837

Satou, Y., Mineta, K., Ogasawara, M., Sasakura, Y., Shoguchi, E., Ueno, K., et al. (2008). Improved genome assembly and evidence-based global gene model set for the chordate *Ciona intestinalis*: new insight into intron and operon populations. *Genome Biol.* 9:R152.

Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., et al. (2002). A cDNA resource from the basal chordate *Ciona intestinalis*. *Genesis* 33, 153–154.

Shi, W., Peyrot, S. M., Munro, E., and Levine, M. (2009). FGF3 in the floor plate directs notochord convergent extension in the *Ciona* tadpole. *Development* 136, 23–28. doi: 10.1242/dev.029157

Shih, J., and Fraser, S. E. (1995). Distribution of tissue progenitors within the shield region of the zebrafish gastrula. *Development* 121, 2755–2765. doi: 10.1242/dev.121.9.2755

Stemple, D. L. (2005). Structure and function of the notochord: an essential organ for chordate development. *Development* 132, 2503–2512. doi: 10.1242/dev.01812

Stöhr, P. (1895). Über die Entwicklung der Hypochorda und des dorsalen Pankreas bei *Rana temporaria*. *Morphol. Jahrb.* 23, 123–141.

Takatori, N., Satou, Y., and Satoh, N. (2002). Expression of hedgehog genes in *Ciona intestinalis* embryos. *Mech. Dev.* 116, 235–238. doi: 10.1016/s0925-4773(02)00150-8

Teillet, M.-A., Lapointe, F., and Le Douarin, N. M. (1998). The relationships between notochord and floor plate in vertebrate development revisited. *Proc. Natl. Acad. Sci. USA* 95, 11738–11738.

Tendeng, C., and Houart, C. (2006). Cloning and embryonic expression of five distinct sfrp genes in the zebrafish *Danio rerio*. *Gene Expr. Patterns* 6, 761–771. doi: 10.1016/j.modgep.2006.01.006

Tessier-Lavigne, M., Placzek, M., Lumsden, A. G. S., Dodd, J., and Jessell, T. M. (1988). Chemotropic guidance of developing axons in the mammalian central nervous system. *Nature* 336, 775–778. doi: 10.1038/336775a0

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., et al. (2001). *Expression of the Zebrafish Genome During Embryogenesis. (NIH RO1 RR15402-01). ZFIN Direct Data Submission*. Available online at: https://zfin.org/ZDB-PUB-010810-1. (accessed May 21, 2021).

Yan, Y. L., Hatta, K., Riggleman, B., and Postlethwait, J. H. (1995). Expression of a type II collagen gene in the zebrafish embryonic axis. *Dev. Dyn.* 203, 363–376. doi: 10.1002/aja.1002030308

Yasuo, H., and Satoh, N. (1993). Function of vertebrate T gene. *Nature* 364, 582–583. doi: 10.1038/364582b0

Yokomori, R., Shimai, K., Nishitsuji, K., Suzuki, Y., Kusakabe, T. G., and Nakai, K. (2016). Genome-wide identification and characterization of transcription start sites and promoters in the tunicate *Ciona intestinalis*. *Genome Res.* 26, 140–150. doi: 10.1101/gr.184648.114

Conflict of Interest: 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 © 2021 Oonuma, Yamamoto, Moritsugu, Okawa, Mukai, Sotani, Tsunemi, Sugimoto, Nakagome, Hasegawa, Shimai, Horie and Kusakabe. 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) and the copyright owner(s) 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 Ecology and Evolution

REVIEW
published: 29 June 2021
doi: 10.3389/fevo.2021.703163

Improved Understanding of the Role of Gene and Genome Duplications in Chordate Evolution With New Genome and Transcriptome Sequences

Madeleine E. Aase-Remedios and David E. K. Ferrier*

The Scottish Oceans Institute, Gatty Marine Laboratory, School of Biology, University of St Andrews, St Andrews, United Kingdom

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Ricard Albalat,
University of Barcelona, Spain
Ignacio Maeso,
Andalusian Center for Development Biology (CABD), Spain

****Correspondence:***
David E. K. Ferrier
dekf@st-andrews.ac.uk

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Ecology and Evolution

Received: *30 April 2021*
Accepted: *08 June 2021*
Published: *29 June 2021*

Citation:
Aase-Remedios ME and Ferrier DEK (2021) Improved Understanding of the Role of Gene and Genome Duplications in Chordate Evolution With New Genome and Transcriptome Sequences. Front. Ecol. Evol. 9:703163. doi: 10.3389/fevo.2021.703163

Comparative approaches to understanding chordate genomes have uncovered a significant role for gene duplications, including whole genome duplications (WGDs), giving rise to and expanding gene families. In developmental biology, gene families created and expanded by both tandem and WGDs are paramount. These genes, often involved in transcription and signalling, are candidates for underpinning major evolutionary transitions because they are particularly prone to retention and subfunctionalisation, neofunctionalisation, or specialisation following duplication. Under the subfunctionalisation model, duplication lays the foundation for the diversification of paralogues, especially in the context of gene regulation. Tandemly duplicated paralogues reside in the same regulatory environment, which may constrain them and result in a gene cluster with closely linked but subtly different expression patterns and functions. Ohnologues (WGD paralogues) often diversify by partitioning their expression domains between retained paralogues, amidst the many changes in the genome during rediploidisation, including chromosomal rearrangements and extensive gene losses. The patterns of these retentions and losses are still not fully understood, nor is the full extent of the impact of gene duplication on chordate evolution. The growing number of sequencing projects, genomic resources, transcriptomics, and improvements to genome assemblies for diverse chordates from non-model and under-sampled lineages like the coelacanth, as well as key lineages, such as amphioxus and lamprey, has allowed more informative comparisons within developmental gene families as well as revealing the extent of conserved synteny across whole genomes. This influx of data provides the tools necessary for phylogenetically informed comparative genomics, which will bring us closer to understanding the evolution of chordate body plan diversity and the changes underpinning the origin and diversification of vertebrates.

Keywords: amphioxus, tunicates, lamprey, ohnologues, subfunctionalisation, rediploidization, allopolyploidy, autopolyploidy

INTRODUCTION

The genetic basis of development is of fundamental interest to understanding animal evolution. In the course of the evolution of chordates, gene and genome duplications have played important roles in shaping gene families by providing the means with which genetic novelties can arise (Ohno, 1970). Whole genome duplications (WGDs) have occurred at the base of speciose and diverse lineages, including the vertebrate two rounds of whole genome duplication (2R WGD) (Meyer and Schartl, 1999; Dehal and Boore, 2005), or the teleost-specific 3R WGD (Hoegg et al., 2004; Meyer and Van de Peer, 2005), and have been implicated in the evolution of more complex, novel, or diverse clades. The 2R WGD duplicated the genome of the vertebrate ancestor twice, such that many genes were retained in as many as four copies, though the majority of paralogues were lost (Dehal and Boore, 2005). The paralogues that were retained in duplicate, ohnologues, are of particular importance to understanding chordate evolution because they may be candidate genes underpinning the evolutionary transition to the vertebrates, and many of these genes retained in duplicate have roles in development (Brunet et al., 2006). While the 2R WGD certainly affected the evolution and subsequent radiation of the vertebrates, duplications also occur frequently on much smaller scales than polyploidisation, and small-scale duplications of developmental genes have also played a role in chordate evolution. Our current understanding of how gene and genome duplications have shaped chordate lineages has progressed rapidly due to the advancements in comparative genomics and with the availability of new genome and transcriptome sequences from phylogenetically informative lineages like amphioxus, lampreys, coelacanth, and monotremes.

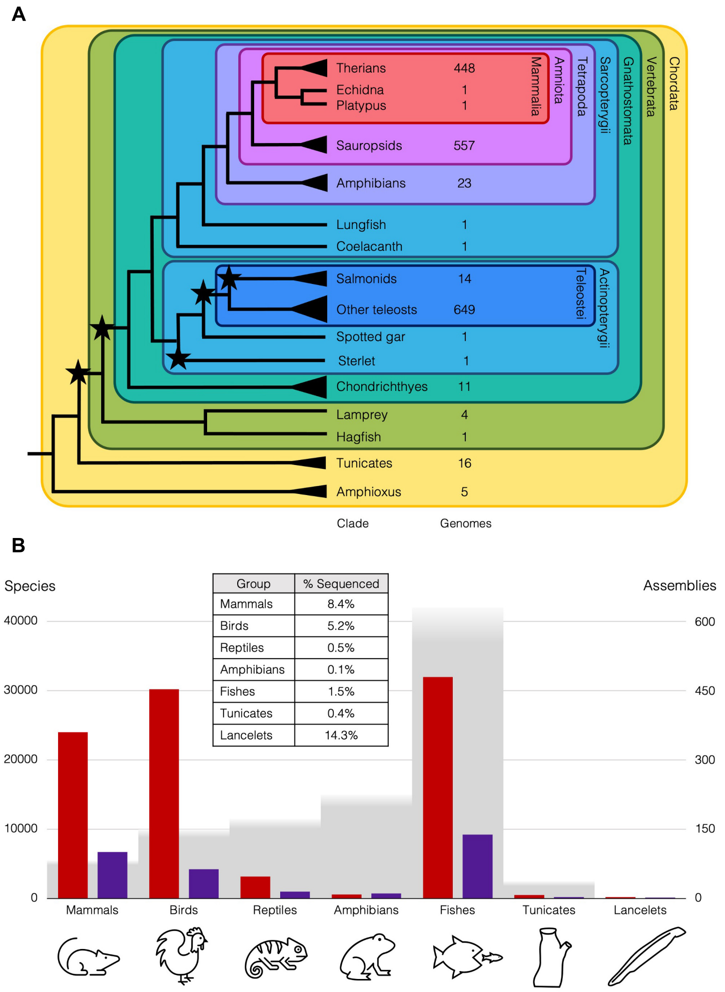
ADDRESSING THE ORIGIN OF VERTEBRATES

Within the chordates, there are three extant subphyla, the cephalochordates, the tunicates/urochordates, and vertebrates, the vertebrates being the most familiar and speciose of the three, which in turn includes cyclostomes (lampreys and hagfish), actinopterygians (ray-finned fish), amphibians, sauropsids (birds and reptiles), and mammals (**Figure 1**). The three chordate lineages share common traits including a hollow neural tube located dorsal to a notochord, pharyngeal slits, and a post-anal tail, but the vertebrate lineage is far more speciose and diverse than either invertebrate chordate subphylum. The field of developmental biology has spent decades determining the genetic basis for the construction of the morphologies of a select few model species, including a handful of vertebrates (like mouse, frog, chick, and zebrafish), but in our current age of genomics, our understanding is growing at an unprecedented pace, with much broader taxon-sampling now making possible a more "phylogenetically informed" and hence evolutionarily robust era of comparative biology and evolutionary developmental biology.

Invertebrate Chordate Models

One of the major questions facing evolutionary developmental biology is the origin of the vertebrates. For a comparison, sister to the vertebrates is the tunicate lineage, the two lineages together making up the clade Olfactores (Delsuc et al., 2006). Tunicates generally have two life stages, comprising the tadpole larval stage and the sessile filter-feeding adult in which chordate-like larval traits such as the notochord, tail, and head are absent or highly reorganised (Lemaire, 2011). Tunicate models such as *Ciona* spp. have been used for decades, representing a growing basis of comparative genomics and transcriptomics studies in chordate developmental biology. Tunicate embryos have been used as a model for the chordate nervous system, as well as for key chordate traits such as the notochord and the precursors of the vertebrate neural crest cells (Schubert et al., 2006; Kugler et al., 2011; Holland, 2016). Neural crest cells are characteristic of vertebrates and constitute a population which migrates and differentiates to contribute to the vertebrate peripheral nervous system, pigmented cells, and elements of the skeletal system, thus the origin of these cells is of particular importance to understanding the invertebrate–vertebrate transition (Gans and Northcutt, 1983). While invertebrate chordates are considered to lack a true neural crest, they do have cells with some similar properties. These include migratory cell types emerging from the neural plate boundary and differentiating into neurons or pigment cells but, critically, the skeletogenic role is lacking from these tunicate cells (Jeffery et al., 2004; Jeffery, 2006; Stolfi et al., 2015). Tunicates have also been instrumental in understanding the control of gene expression because of their relatively simple body plan and a small number of transcription factors compared to vertebrate models, and they have greater extents of homology to vertebrates than other comparisons outside of the phylum Chordata (Lemaire, 2011). There is a growing database of genomic resources for ascidians (ANISEED), which in recent years has acquired datasets focussing on gene expression, including RNA-seq and epigenetic datasets, as well as many new genome annotations (Brozovic et al., 2018), reflecting the utility of tunicate models for molecular biology. The most recent update to this database includes the addition of a larvacean, *Oikopleura dioica*, and the expansion of tools to explore gene expression data in larvae (Dardaillon et al., 2019).

Tunicate genomes, while small, evolve very quickly and have undergone extensive rearrangements relative to the other chordate subphyla (Denoeud et al., 2010; Holland, 2016). At first, this made genomic comparisons between tunicates and vertebrates confusing, and more than a fifth of *Ciona robusta* (previously *C. intestinalis* type A) genes had no clear homology to known bilaterian genes when the genome was first sequenced (Cañestro et al., 2003). Gene families typically conserved in clusters in the genomes of other organisms are often disorganised in tunicate genomes, e.g., the Hox cluster (Sekigami et al., 2017). There is also an issue with long-branch attraction in molecular phylogenies with the highly divergent *C. robusta* sequences, but the sequencing of a second clade of tunicates, the larvacean *O. dioica* to complement the ascidian *C. robusta*, helped resolve this (Delsuc et al., 2006). Now genomes are available for



Group	% Sequenced
Mammals	8.4%
Birds	5.2%
Reptiles	0.5%
Amphibians	0.1%
Fishes	1.5%
Tunicates	0.4%
Lancelets	14.3%

FIGURE 1 | Diversity of chordates and available genome assemblies. **(A)** Cladogram of chordate lineages and count of representative genome assemblies available on NCBI (https://www.ncbi.nlm.nih.gov/assembly/organism; Accessed 25 April 2021). Branches with single species mentioned in the text end with a line, but lineages with multiple species end with a triangle, e.g., there are ~35 amphioxus species (Bertrand and Escriva, 2011). Stars represent WGDs. NCBI Taxonomic codes used to count genomes for each lineage: Therians (taxid:32525); Sauropsids (taxid:8457); Amphibians (taxid:8292); Salmonids (taxid:8015); Teleosts (taxid:32443); Cartilaginous fishes (taxid:7777); Tunicates (taxid:7712); Cephalochordata (taxid:7735). **(B)** Genome assemblies at scaffold (red) and chromosome (purple) level available on NCBI for animal subgroups (https://www.ncbi.nlm.nih.gov/genome/browse#!/eukaryotes/; Accessed 25th April 2021), right axis, relative to approximate number of species (grey), left axis. Table inset is approximate percentage of species in each group that has been sequenced at least to the draft level.

several different tunicate lineages, including a chromosomal level assembly of *O. dioica* (Bliznina et al., 2021). Sequencing of two *C. intestinalis* Type B (i.e., *C. intestinalis* rather than *C. robusta*) genomes revealed that between and even within *Ciona* species, there have been many chromosomal inversions and there are high levels of polymorphism (Satou et al., 2021). Because of their fast rates of evolution, between tunicate and vertebrate genomes there is little conserved synteny remaining, making inferences about the genomic changes at the origin of the vertebrates difficult from these comparisons.

2R and Amphioxus

It is now widely understood that 2R WGD occurred in the ancestors leading to vertebrates, and these events have been implicated as playing a permissive role in the evolution of vertebrate novelties. Still, the specifics of how the 2R WGD provided the genetic basis for these traits is not yet fully understood. Much of the evidence comes from comparative studies with amphioxus, the basal chordate lineage. Considered the "archetypal" chordate, both its genome and morphology are thought to have not changed much in the nearly 600 million years since it diverged, in contrast to the fast-evolving tunicates (Holland et al., 2008; Putnam et al., 2008; Zhang et al., 2018), making it an ideal model for inferring features of the chordate ancestor (Dehal and Boore, 2005; Schubert et al., 2006; Bertrand and Escriva, 2011). While amphioxus lacks a true neural crest, as do tunicates (Holland and Holland, 2001), the draft genome of *Branchiostoma floridae* contains most pro-orthologues of the genes that constitute the regulatory network required for vertebrate neural crest formation (Yu et al., 2008). This suggests that the ancestral chordate may have possessed the components of a neural crest rudiment, but it was the 2R WGD at the origin of the vertebrates that may have facilitated the evolution of the true neural crest.

With the sequencing of an amphioxus genome came the most solid evidence for the 2R WGD, revealing the prevalent one-to-four syntenic relationship between the scaffolds of *B. floridae* and the human genome (Putnam et al., 2008). This amphioxus genome was highly polymorphic, thus each haplotype assembled separately, which complicated these analyses. More recently, the application of newer techniques like long-read sequencing and Hi-C has allowed the genome to be assembled to the chromosome level, revealing the depth of synteny conserved amongst chordate lineages (Simakov et al., 2020). Advancements in long-read sequencing reduce the assembly errors that can arise in highly polymorphic genomes that can cause false gene gains to be inferred for alleles rather than paralogues (Denton et al., 2014). With this chromosome-level assembly, the authors were able to precisely determine the sequence and timing of polyploidisations, and the chromosomal rearrangements that occurred after each one in the process of rediploidisation characterised by chromosomal fusions and rampant gene losses (Simakov et al., 2020).

Cyclostomes—the Agnathan Lampreys and Hagfish

This recent study also took a comparative approach to determine the timing of 2R WGD in vertebrate ancestors, with particular focus on the placement of the divergence of the cyclostome lineage. The cyclostomes (lampreys and hagfish) diverged from the gnathostomes 500 Mya, and provide a phylogenetically well-placed point of comparison to the jawed vertebrates (Janvier, 2007). They share a cartilaginous skeleton including a cranium, paired sensory organs, and true neural crest, but lack key gnathostome structures like a jaw and paired fins (Green and Bronner, 2014). An initial lamprey genome assembly from a sample of liver tissue was later updated with a germline assembly, since hundreds of genes are eliminated from somatic tissues, though both assemblies revealed a high number of rearrangements and repeats (Smith et al., 2013, 2018). These complicated genomic characteristics have hindered the genome assembly and contributed to uncertainty about the number of duplications in the cyclostome lineage, and how many, if any, are shared with the gnathostomes.

The Timing of 2R

One potential marker of genome duplications is the Hox cluster. This cluster of developmental transcription factor genes is highly conserved across animals, so that comparison of the single cluster in amphioxus to the four in mammals gave an early line of evidence for 2R (Garcia-Fernàndez and Holland, 1994) and was complemented by the discovery of the teleost-specific 3R and the seven (i.e., eight minus one) Hox clusters in zebrafish (Amores, 1998). Lampreys were reported to have at least six Hox clusters, and the updated germline assembly of *Petromyzon marinus* revealed the synteny of these regions shared with other vertebrates' Hox chromosomes (Mehta et al., 2013; Smith et al., 2018). Amongst these six, two pairs appear to be the product of more recent, chromosome-level duplications, suggesting there may be a high prevalence of large-scale duplications occurring more recently besides the WGDs in lampreys. Smith and Keinath (2015) had previously constructed the lamprey meiotic map using RAD-seq, and with these syntenic comparisons, concluded that fewer chromosomal rearrangements were required if lampreys diverged after 1R WGD, but before 2R.

These data are intriguing, and indeed several different scenarios of duplications could result in the six Hox clusters, among other unexpected or unique patterns of synteny. Lampreys could have undergone the 2R WGD and several independent chromosomal duplications and losses or diverged after 1R then undergone their own WGD or several chromosome-level duplications (**Figure 2**). If lampreys and gnathostomes share 2R, the lack of direct orthology between lamprey Hox clusters and those of gnathostomes can be explained by cyclostome divergence during rediploidisation, before the ancestral gnathostome karyotype was established. This pattern, where orthology is obscured because lineages retained different paralogues when they diverged, termed tetralogy, has been described for early-branching 3R teleosts, and may apply here as well (Martin and Holland, 2014). Many of the paralogues in lamprey are younger than 2R ohnologues, suggesting some independent duplications have occurred more recently than either of the 2R WGD (Mehta et al., 2013). Some phylogenies suggest that cyclostomes and gnathostomes share both WGDs (Kuraku et al., 2009), while others support independent duplications (Force et al., 2002) or that cyclostomes diverged before either WGD (Fried et al., 2003), though these early conclusions may have been misled by long-branch attraction and hidden paralogy due to the divergent sequences and independent gene gains and losses in cyclostome genomes. Furthermore, with only one lamprey genome assembled as of 2018, the rigour of any of these conclusions was tenuous (Holland and Ocampo Daza, 2018). Because of this, the extent to which lamprey duplications may be shared with the other cyclostome lineage or if one or both of the 2R WGD occurred in the vertebrate ancestor has been contentious. Previous synteny evidence suggested that all vertebrates share at least 1R, and likely the 2R as well (Holland and Ocampo Daza, 2018; Sacerdot et al., 2018). However, with the new chromosome-level assembly of *B. floridae*, the distinct chromosomal fusion events that occurred between the first and second round of WGD could be discerned. Simakov et al. (2020) provide evidence that lampreys diverged after only 1R, as certain genomic rearrangements predating the 2R are unique to jawed vertebrates and lacking from lampreys, in particular the gnathostome-specific CLGE-CLGO fusion, and the observation that lamprey syntenic blocks are generally unmixed. This new assembly improves upon the extensive synteny work aimed at determining the origin of chordate karyotypes from previous studies (Nakatani et al., 2007; Putnam et al., 2008; Sacerdot et al., 2018) and provides the latest working hypothesis for the

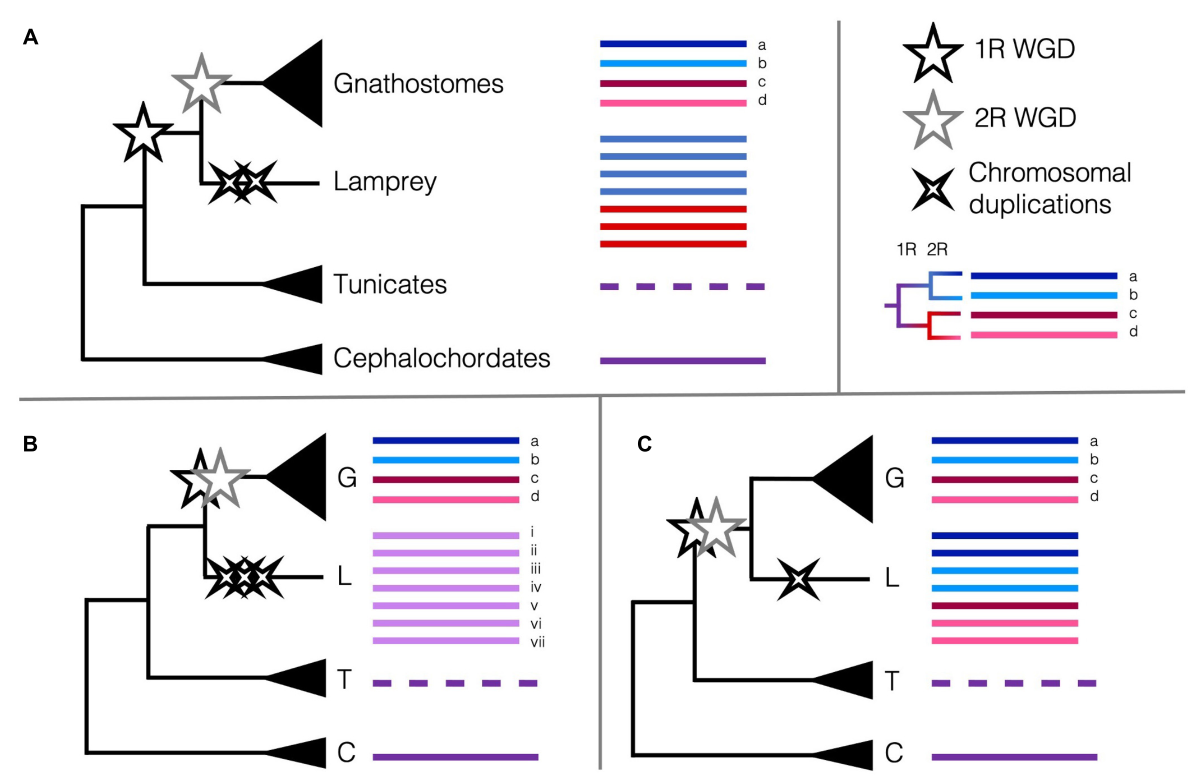


FIGURE 2 | Scenarios of the timing of the 2R WGD relative to the divergence of cyclostomes and the Hox cluster as a marker of WGDs. Pre-duplicate chordate phyla with single Hox clusters are coloured to represent the ancestral state (purple lines). The reorganised Hox cluster in tunicates is represented by a dashed line. **(A)** The current consensus that cyclostomes diverged from the gnathostomes after 1R (black star) but before 2R (grey star) and underwent some independent chromosome-level duplications (black crosses) (Smith et al., 2018; Simakov et al., 2020). Remnants of 1R paralogy (red versus blue) may distinguish Hox clusters in the cyclostomes. **(B)** Alternate hypothesis that cyclostomes diverged before 1R and only underwent several chromosome-level duplications (Fried et al., 2003; Furlong et al., 2007) (or perhaps an independent WGD). **(C)** Alternate hypothesis that cyclostomes diverged after 2R and underwent few independent chromosome-level duplications (Caputo Barucchi et al., 2013; Sacerdot et al., 2018). Remnants of 1R paralogy and 2R paralogy (dark red versus pink, dark blue versus light blue) may distinguish Hox clusters in the cyclostomes. Cyclostome Hox cluster numbering and colouring does not represent actual homology [(Smith et al., 2018) found homology between lamprey Hox clusters i. and iv., and ii. and v.]; no conclusive affinities have been found between cyclostome Hox clusters and any particular gnathostome Hox cluster.

timing of the 2R WGD and their impact on the evolution of vertebrates.

Cyclostome Peculiarities

Genomes of amphioxus and lamprey have been instrumental in understanding the origin of the vertebrates as these lineages are phylogenetic outgroups to larger chordate lineages. While amphioxus is remarkably conserved, making inferences about the chordate ancestor fairly straightforward, the lamprey genome is much more complicated. These complications, however, are interesting themselves, especially the programmed genomic rearrangements (PGRs) in the somatic tissues, presumably an adaptation reflecting the genetic conflict between germline and soma (Smith et al., 2012; Timoshevskiy et al., 2017). The hagfish genome also undergoes PGR (Nakai et al., 1991, 1995), and further resolution of genomes in this cyclostome clade should be instrumental in deciphering the mechanism of PGR. In addition, comparisons to determine homology of structures and their genetic underpinnings depend on reliable gene annotations. For instance, while amphioxus lacks myelin, lampreys were reported to have it (Smith et al., 2013), though lamprey neurons are not myelinated. In fact, the genes annotated as myelin in lamprey are likely to be in another gene family altogether (Werner, 2013). This highlights the pitfalls of automated gene annotations and the need for careful gene curation and annotation with a focus on phylogenetically informed and robust gene nomenclature.

More recently, the genome sequence of a Reissner's lamprey (*Lethenteron reissneri*) was assembled to higher completeness than any previous lamprey assemblies, aided by the combinatorial use of short- and long-read sequencing, and Hi-C (Zhu et al., 2021). The protein-coding genes of this genome were annotated using a wider and more phylogenetically informed array of species, including several genomes only recently sequenced from both invertebrates and vertebrates, revealing a total of seven Hox clusters (**Figure 2**). This new assembly, as well as the improved sea lamprey genome (Smith et al., 2018) show that the independent duplications indicated by the seven Hox clusters are likely synapomorphic to at least the lampreys, but maybe to the cyclostomes as a whole. Comparisons of lamprey and hagfish Hox clusters revealed that cyclostomes share the loss of Hox genes from paralogy group 12 and likely independent duplications, as hagfish may have six or more Hox clusters as well (Pascual-Anaya et al., 2018). These data are taking us closer to finally determining the cyclostomes' placement in relation to the 2R WGD. The sequencing of multiple cyclostomes, particularly from hagfishes, would build on current data in order to enable further understanding of the role 2R WGD played in chordate evolution, and particularly, the evolution of vertebrate- or gnathostome-specific developmental novelties.

DUPLICATIONS IN THE HOX CLUSTER

As mentioned above, the Hox gene cluster has been a fruitful system for determining how gene duplication has affected animal development. Not only have multiple ohnologous clusters been maintained following WGDs, including 2R as well as the teleost-specific 3R and others (Hoegg and Meyer, 2005), but the cluster (and others) arose via extensive tandem duplications of a single ANTP-class transcription factor gene early in animal evolution (Butts et al., 2008; Holland, 2013; Ferrier, 2016). Furthermore, there tends to be a link between the expression of the Hox genes and their clustered arrangement, as genes at one end of the cluster are expressed anteriorly, and expression follows successively toward the other end of the cluster, where those genes are expressed in the posterior of the embryo (Duboule and Dollé, 1989; Graham et al., 1989; Duboule, 2007; Tschopp et al., 2009). This spatial collinearity is frequently observed alongside temporal collinearity, where the activation of the genes follows the same sequence of early to late as anterior to posterior in many animals (Duboule, 1994; Kmita and Duboule, 2003; Iimura and Pourquié, 2006; Monteiro and Ferrier, 2006; Gaunt, 2018; Krumlauf, 2018; Ferrier, 2019). Hox genes are transcription factors, and their expression in the developing embryo is strictly regulated so as to provide positional identity along the anterior–posterior axis of the animal. Changes to the expression of Hox genes therefore changes rostro-caudal identity, most obviously manifested in vertebrates in the formation of the axial skeleton (Casaca et al., 2014). Between different vertebrate groups, differential expression of Hox genes drives the different identities of vertebral segments, illustrating how the Hox gene diversification directly underpins the diversity of vertebrate body-plans (Krumlauf, 1994; Burke et al., 1995). The Hox cluster perfectly illustrates the interplay between gene regulation, expression, duplication, and subfunctionalisation, and the effects of these processes on development and animal evolution.

Hox Clusters Through WGDs

Other studies have turned to the Hox cluster in order to understand the impact of the 3R WGD. Early teleost genome projects revealed excess Hox clusters in zebrafish, medaka, and two pufferfish (Hoegg and Meyer, 2005). In these, as above, the genome sequence and transcriptomics of an appropriate outgroup was essential. PCR screening of the birchir, a non-teleost ray-finned fish, revealed only four Hox clusters, which allowed the timing of the 3R WGD to be determined (Ledje et al., 2002). Another more basally branching fish lineage, the spotted gar, also has four Hox clusters (Braasch et al., 2016) while zebrafish has seven (Amores, 1998). As many as eight Hox clusters can be found in different teleost lineages, though some have been lost from different lineages (Sato and Nishida, 2010; Martin and Holland, 2017). Detection of the 3R WGD has led to speculation that WGDs may play a permissive role in the evolution of diverse lineages, as within the more than 30,000 actinopterygian species, only 50 are non-teleosts which did not undergo 3R (Near et al., 2012). The link of WGDs and subsequent species diversity has been posited for vertebrates as well (Cañestro et al., 2013), but evidence for this link is still lacking (Glasauer and Neuhauss, 2014).

A more recent WGD allows the study of rediploidisation as it is occurring. Salmonids have thirteen Hox clusters (Mungpakdee et al., 2008), indicative that this lineage has undergone a fourth whole genome duplication (4R WGD) more recently, around 88 Mya (**Figure 1**; Macqueen and Johnston, 2014; Lien et al., 2016). The Atlantic salmon (*Salmo salar*) genome was particularly difficult to assemble because of the repetitiveness created by the fourth more recent WGD, but its assembly enables our understanding of the processes of rediploidisation (Lien et al., 2016). Many of the paralogues created in the 4R WGD were able to be identified, even those that were recently pseudogenised (Lien et al., 2016). In another salmonid, the rainbow trout (*Oncorhynchus mykiss*), nearly half of the 4R ohnologues remain in duplicate (Berthelot et al., 2014). Many of the 4R ohnologue pairs in salmon consist of one gene that maintained the ancestral function, as conserved between the salmon and pike, while the second had a different expression pattern (Lien et al., 2016; Robertson et al., 2017). Genomes of 4R salmonids also enable detection of a category of genes retained following successive WGDs, which is enriched in genes involved in development and transcription factors (Berthelot et al., 2014). This was associated with changes to the ecology of salmonids, but did not directly cause rapid species diversification as was hypothesised for the 3R in teleosts, rather, it is correlated with the adaptation of anadromy (Macqueen and Johnston, 2014). These observations suggest that environmental adaptation may have had a larger impact on teleost diversity than the 3R WGD because there are significant discrepancies between the timing of 3R and the divergence of major teleost lineages (Donoghue and Purnell, 2005; Santini et al., 2009; Near et al., 2012), as well as 40–50 million years between the salmonid 4R WGD and their subsequent radiation (Macqueen and Johnston, 2014). The salmonid clade is still within the process of rediploidisation, and comparisons between lineages reveals that a quarter of the Atlantic salmon genome underwent rediploidisation independently of the trout, suggesting that rediploidisation occurred before and after speciation (Robertson et al., 2017). These differently resolved regions may underpin the differences in ecology amongst salmonid lineages and also help to explain the (perhaps surprising) length of time that is required post-WGD before species diversification might be observable (Robertson et al., 2017). The differences between salmon and trout karyotypes also provide support for tetralogy, the pattern of homology between genes in lineages that diverged during rediploidisation (Martin and Holland, 2014). These assemblies of salmon and trout revealed the processes of rediploidisation and provided a case study for examining the fate of recently duplicated genes in the context of a WGD.

THE FATE OF DUPLICATED GENES

Several mechanisms have been proposed for the maintenance of duplicated genes, despite the likelihood of non-functionalisation of redundant paralogues (Nakatani et al., 2007). For some gene types, increased copy number is advantageous, despite their redundancy, such as ribosomal RNAs (Zhang, 2003). Alternatively, once a dosage-sensitive network of genes is duplicated in entirety, perhaps along with the entire genome following a WGD, to preserve the balance of the network components, no loss can occur, therefore entire duplicate networks are retained (Lynch and Conery, 2000). This dosage compensation theory, however, does not result in as high an incidence of retained paralogues being related to dosage-sensitive complexes and pathways as one might expect; instead, this model might better explain a stop-gap mechanism that prevents initial loss and allows time for sub- or neo-functionalisation processes to occur (Hughes et al., 2007).

While duplicated genes are likely to be redundant initially, this redundancy presumably reduces selective constraint and allows one or both daughter genes to evolve, which explains how duplication can facilitate evolutionary novelties. Under the duplication-degeneration-complementation (DDC) model, genes undergo degenerative mutations following their duplication, resulting in the complementary partitioning of the ancestral gene's functions among the daughters (Force et al., 1999). This mechanism of subfunctionalisation is particularly relevant to developmental genes, which have complex *cis*-regulatory regions, since degenerative mutations between the regulatory regions of the daughter genes can quickly partition their expression profiles and therefore their functions. One step further is neofunctionalisation, which suggests that the reduced constraint on paralogues allows for novel functions to evolve in one or both of the paralogues, including new upstream transcription binding sites or even protein domains [sometimes referred to as the duplication-degeneration-innovation (DDI) model] (Jimenez-Delgado et al., 2009). The role of DDC or DDI in the evolution of developmental genes is supported by the overrepresentation of transcription factors among retained paralogues, the often complicated *cis*-regulatory regions these transcription factors have, and the fact that expression rather than sequence changes are the primary observed differences between different animals' developmental genes.

Evidence of Subfunctionalisation of 3R Paralogues

Evidence supporting the role of DDC, or DDI, in chordate evolution is growing. Large RNA-seq datasets have been instrumental in detecting subfunctionalised ohnologues, for instance the reduced expression of genes retained in duplicate compared to single orthologues in mammals (Qian et al., 2010). One of the largest datasets of expression data of paralogues and their pro-orthologues in another species comes from the gar genome project. The spotted gar is a ray-finned fish whose lineage diverged from the teleosts before the 3R WGD that characterises this latter clade (**Figure 1**; Braasch et al., 2016). This makes it an ideal outgroup to understand the impact of the 3R WGD on teleost evolution that has been implicated in possibly permitting the huge diversity of fishes (Hoegg et al., 2004; Sato and Nishida, 2010). This model suggests that speciation occurs following WGD, but before the genomes undergo rediploidisation, so that changes to genome architecture like chromosomal fusions reinforce speciation (Taylor et al., 2001). Because it has not undergone the 3R, but still lies within Actinopterygii, the spotted gar possesses pro-orthologues of the duplicated 3R ohnologues. Thus, expression of gar genes is expected to be similar to that of the ancestral gene. Between the gar and teleost transcriptomes, each of the two teleost 3R ohnologues were expressed at lower levels than the gar pro-orthologue, but the pooled expression of the ohnologue pair was similar to the gar gene's expression (Braasch et al., 2016). This pattern suggests that following 3R, many teleost ohnologues underwent subfunctionalisation in accordance with the predictions of the DDC model, and for certain gene families, subfunctionalisation amongst paralogues has been traced to specific regulatory regions.

These inter-species comparisons are giving us concrete examples of the DDC process. For example, comparisons between mammals and zebrafish revealed that in the gene-poor region around the *Pax6* locus, there are cross-functional conserved non-coding elements (CNEs) that have been retained in a patchwork pattern between the zebrafish *Pax6* paralogues relative to the single mouse *Pax6* (Navratilova et al., 2009). The zebrafish, a 3R teleost, has two *Pax6* paralogues: *pax6a* and *pax6b*, which are expressed in overlapping yet complementarily constricted patterns relative to tetrapod *Pax6*; *pax6a* is expressed widely in the eye and brain, while *pax6b* is found in the eye, a restricted section of the developing brain, and the pancreas (Kleinjan et al., 2008). The changes to expression of the two ohnologues was traced to the differential loss of certain *cis*-regulatory regions around the two genes; brain-specific elements from *pax6b* and the pancreas control element from *pax6a* (Kleinjan et al., 2008). This pattern of retention of the two *Pax6* ohnologues in zebrafish is consistent with the DDC framework of subfunctionalisation, illustrating the role of *cis*-regulatory elements in this process.

Specialisation of Paralogues

Among daughter paralogues, subfunctionalisation processes are not necessarily symmetrical. Specialisation results in one paralogue with a function more similar to that of the ancestral gene, while the other paralogue's function diverges, either to a small subset of the ancestral function, or even the adaptation of new functions (Farrè and Albà, 2010; Marlétaz et al., 2018). Specialisation of ohnologues has been detected often in the past but has been described by several different mechanisms, corresponding to the different mechanisms of paralogue retention following duplication. Previous studies focusing on protein sequence evolution and detecting neofunctionalisation have found examples of specialisation in protein function (Chain and Evans, 2006; Steinke et al., 2006; Sémon and Wolfe, 2008). This takes place in accordance with the process described as qualitative subfunctionalisation, driven by adaptive evolution of the paralogue's protein sequence (Espinosa-Cantú et al., 2015). For many developmental genes, however, the DDC hypothesis, which falls under quantitative subfunctionalisation and occurs by neutral processes affecting gene regulation, is more relevant (Espinosa-Cantú et al., 2015; Braasch et al., 2018). Following duplication, for some gene families, one paralogue retains widespread expression across several domains consistent with the inferred ancestral function, while the other's function is reduced to a specialised domain. This process is linked to the changes to regulatory domains occurring primarily on one paralogue instead of symmetrically across both (during the degeneration and complementation aspects of the DDC), though this was detected for the most part in genes with complex or widespread ancestral expression patterns (Marlétaz et al., 2018). Some studies of expressed sequence tag (EST) data between the polyploid frog *Xenopus laevis* and the non-duplicate species *X. tropicalis* found that many paralogues exhibited subfunctionalisation, but others had patterns where only one of the two paralogues was greatly reduced in expression (Hellsten et al., 2007). Also, for many pairs of 4R ohnologues in the salmon, one retained a broader more ancestral expression pattern, while only the other ohnologue changed, often even neofunctionalising (Lien et al., 2016). These observations of specialisation are consistent with an asymmetric mechanism of the DDC, where one ohnologue's regulatory region degenerates much more than the other.

A more recent detection of this pattern extends more broadly across the amphioxus transcriptome, with comparisons to multiple vertebrates (Marlétaz et al., 2018). For certain ohnologues, new regulatory elements were found near the specialised gene, indicating a role for neofunctionalisation as well as the asymmetrical loss of one ohnologue's "ancestral" regulatory elements (Marlétaz et al., 2018). These studies rely on broad overviews of gene expression from extensive transcriptome datasets, so may be unable to detect more focused and specific changes to gene function and expression between particular ohnologues, such as qualitative changes to protein sequences described above. Furthermore, specialisation may only be detectable for genes whose ancestral function or expression was widespread and complex (pleiotropic), while other studies detecting asymmetrical changes in expression amongst paralogues with more specialised or localised functions ascribe the process to sub- or neo-functionalisation. The mechanisms resulting in more symmetrical patterns of subfunctionalisation, namely the relaxed constraint provided by redundancy allowing degenerative mutations to accrue in regulatory regions, are not required to affect each paralogue equally, so patterns described as specialisation or neofunctionalisation can also occur by the same logic, if for some reason one paralogue is affected more than the other. This asymmetry amongst paralogues can also explain how novelties arose via co-option of redundant ohnologues to novel functions, through changes in gene regulation.

Regulatory Elements and Subfunctionalisation

Advancements in bioinformatic approaches to comparative genomics have facilitated the detection of CNEs, aiding in the identification of the regulatory changes between paralogues. The patterns of retention of these CNEs, many of which are situated within the introns of neighbouring genes, may have important consequences for the evolutionary maintenance of the synteny of gene clusters or highly conserved genomic neighbourhoods, in which many developmental genes can be found. For example, VISTA plots show that the conserved regulatory elements for *Pax6* are situated in the introns of several gene neighbours across vertebrates (Kleinjan et al., 2008; Navratilova et al., 2009), providing the selective pressure to maintain these gene clusters. This can be seen in many other developmental gene clusters, including the MyoD-related genes that are key for muscle development. In most non-teleost 2R vertebrates, there are four myogenic regulatory factors (MRFs; *MyoD*, *Myog*, *Myf5*, and *Myf6*), two of which, *Myf5* and *Myf6*, reside in a closely linked cluster and share many overlapping regulatory elements (Braun et al., 1990; Carvajal et al., 2008). While in zebrafish these genes act early in myoblast determination and later in myoblast differentiation, respectively, so that *Myf5* expression is followed by that of its neighbour, *Myf6*, in mice *Myf6* also has an early role in determination and so has a biphasic expression pattern (Kassar-Duchossoy et al., 2004; Schnapp et al., 2009; Moncaut et al., 2013; Hernández-Hernández et al., 2017). This early versus late distinction is exemplified by *MyoD* binding the *Myog* promoter (Tapscott, 2005). Deletion of the minimal *Myf5* promoter results in misexpression of *Myf6* in *Myf5* regions, thus it is clear that the regulatory domains of these two genes are intrinsically linked (Carvajal et al., 2008).

Mirroring this cluster, a cryptic fifth MRF ohnologue is located adjacent to *MyoD* in the genomes of the coelacanth, sterlet, and spotted gar (Aase-Remedios et al., 2020). Only with the recent availability of the genomes of these "non-model" organisms could this fifth ohnologue, *Myf7*, be found, since it has been lost from most other vertebrate lineages to which common model organisms belong, including tetrapods, cartilaginous fish, and teleosts. The finding of this fifth vertebrate MRF upended previous interpretations of the evolution of the four MRFs shared amongst all vertebrates. Instead of a single ancestral gene, duplicated in 2R into four ohnologues, and a highly unusual translocation event, it is much more likely that there was an ancestral two-gene state resulting from a tandem duplication that generated the two types of MRF, early and late (Aase-Remedios et al., 2020). This cluster nevertheless shares synteny with the *MyoD* locus across the vertebrates, potentially as a result of constraint on regulatory elements that overlap with the MRFs and their gene neighbours. Other developmental gene families also have so-called "cryptic" vertebrate paralogues analogous to *Myf7*, that have been lost from common model system lineages and found only with the availability of genomes from more basally branching species (Kuraku et al., 2016), such as *Bmp16* (Feiner et al., 2019) and *Foxl2B* (Geraldo et al., 2013). The wider availability of sequencing has allowed more phylogenetically informative outgroups and lineages to be studied, changing our understanding of chordate evolution.

INDEPENDENT DUPLICATIONS

Small-Scale Independent Duplications in Chordate Lineages

Wider taxon-sampling has enabled the detection of independent duplications in certain lineages overlaid on the larger general trend of various WGDs. In many important developmental gene families, amphioxus has undergone its own expansions of gene clusters, formed by tandem duplications (Minguillón et al., 2002). This ranges from expansions within existing clusters, possibly like the Posterior Hox gene *AmphiHox15* within the Hox cluster (Holland et al., 2008), to the advent of new clusters, like the expansion of the amphioxus MRFs (Schubert et al., 2003; Urano et al., 2003; Yuan et al., 2003; Somorjai et al., 2008; Bertrand et al., 2011; Tan et al., 2014; Aase-Remedios et al., 2020). The five genes in this cluster are expressed in different temporal and spatial patterns during amphioxus muscle development, indicating these genes have subfunctionalised as well, providing an independent case of MRF subfunctionalisation from that seen with the vertebrate MRFs (Aase-Remedios et al., 2020). In the *Pax* gene family of developmental transcription factors, we see a comparable case of independent subfunctionalisation between amphioxus and vertebrates. Amphioxus has, independently from vertebrates, duplicated the ancestral chordate pro-orthologue of *Pax3/7* to produce two *Pax3/7 trans*-homologues of vertebrate *Pax3* and *Pax7*, which have again evolved distinct expression patterns (Barton-Owen et al., 2018). Being that in vertebrates, *Pax3* and *Pax7* are upstream regulators of the MRFs, perhaps this duplication is acting in concert with the MRF expansion, as an expanding network of muscle regulators. These observations complicate the overall relationship between the pre-duplicate amphioxus genome and vertebrate genomes shaped by 2R WGD and require careful "mining" of the latest genome sequence data and close consideration of gene phylogenies. Overall, these cases show that although amphioxus does indeed provide us with an excellent proxy for much of the chordate ancestor, this animal also has its own unique evolutionary history. This needs to be carefully considered in order to properly infer the features of the chordate ancestor.

Lineage-Specific WGDs

Wider taxon-sampling allows us to study not only phylogenetically informative lineages, but also species that serve as case studies for interesting phenomena like more recent lineage-specific WGDs. Recent sequencing efforts have provided a characterisation of the WGD in the sterlet lineage (Cheng et al., 2019; Du et al., 2020). Not only does this genome provide an outgroup comparison to 3R teleosts as a basal osteichthyan (**Figure 1**), but because sturgeons are slow-evolving, they have retained ancestral characteristics possibly remnant of the bony fish ancestor. Evidence of several levels of polyploidy has been observed in different sturgeon groups (Vasil'yev et al., 1980; Fontana et al., 2008), but the WGD identified at the base of the sturgeon lineage provides a case study to examine the process of rediploidisation. In the sturgeon, while large chromosomes were retained in duplicate, many of the second copies of the smaller chromosomes were lost, meaning some chromosomes are tetraploid, while others are diploid (Romanenko et al., 2015; Du et al., 2020). In contrast, in 4R salmonids, ohnologues were lost gene by gene (Lien et al., 2016). The sterlet genome's mobilome shows identical transposable element content between each haplotype, suggesting an origin by autopolyploidy (Du et al., 2020). Comparisons to allopolyploid genomes could reveal the different mechanisms of genomic rediploidisation following the different modes of duplication; perhaps an autopolyploid genome allows the loss of entire chromosomes, while allopolyploid genomes are constrained by differences between haplotypes, and therefore losses occur on smaller scales.

An important allopolyploid is the African clawed frog (*Xenopus laevis*) (Kobel and Du Pasquier, 1986; Session et al., 2016). This frog has a diploid chromosome number of 36, almost twice the number of its congeneric *X. tropicalis*, with 20, and arose via hybridisation of two 18n frogs that had undergone chromosomal fusion between homologues of *X. tropicalis* chromosomes 9 and 10 (Evans, 2008). There is extensive conserved homology between the two sub-genomes in *X. laevis* and the diploid *X. tropicalis*, reflecting the recentness of this duplication, and only 17% of duplicated genes have been lost (Uno et al., 2013), in contrast to ∼80% of teleost 3R ohnologues for example (Brunet et al., 2006). The mechanism for retention in the frog may be dosage-based, though some paralogues may also have subfunctionalised (Charbonnier et al., 2002; Hellsten et al., 2007; Sémon and Wolfe, 2008), and different types of genes are lost at different rates (Session et al., 2016). Notably, the karyotype is stable, unlike in the sterlet, probably due to the lack of recombination between the two *X. laevis* sub-genomes (S and L), though more genes are lost from sub-genome S, which has also undergone more intra-chromosomal rearrangements (Session et al., 2016). Altogether, this sheds light on the patterns of rediploidisation in an allopolyploid, where each sub-genome is more intact and independent, as opposed to the sterlet example of an autopolyploid, where the polyploid genome is truly redundant and entire chromosomes can be lost.

These differences in rediploidisation are relevant more widely, as the 2R WGD at the base of the vertebrates is thought to have taken place with an initial autopolyploidy, then subsequent allopolyploidy of two descendants of that first polyploid (Simakov et al., 2020). This is supported by the genomic rearrangements and symmetrical gene losses that follow 1R, also seen in the rediploidisation process in the sterlet. These rearrangements are shared with lampreys, but the asymmetric gene losses and independence of haplotypes characteristic of allopolyploids are not. Ancient duplications can now be understood in more detail with our better understanding of the different rediploidisation processes, only now possible with the wider taxon-sampling of case-study species with lineage-specific WGDs. For instance, only 0.2% of the diversity of amphibian genomes has been sequenced, while over 8% of mammal genomes have been sequenced, and roughly the same number of genomes are available for teleosts as sauropsids, though there are twice as many teleost species (**Figure 1B**). This has changed remarkably over the past few decades with the rapid improvements in genome sequencing, and is set to improve still further with initiatives like the Earth Biogenome Project (Lewin et al., 2018), which should release us from our anthropocentric concentration on only biomedically and economically relevant species and allow us instead to obtain an unbiased overview of genomic biodiversity and the processes that generated it.

NEW GENOMES AND TRANSITIONS WITHIN VERTEBRATES

The Origin of Tetrapods

Of these more recently sequenced species, the coelacanth genome has been particularly instrumental in addressing the origin of tetrapods and the transition to land, as it is the most basal-branching sarcopterygian, to which tetrapods also belong (**Figure 1**). Likewise, the lungfish genome has been instrumental in revealing the genomic changes underpinning this transition, as it is the closest living relative to tetrapods (**Figure 1**; Meyer et al., 2021). These groups make more suitable comparisons to tetrapods than commonly used teleost outgroups, since they have not undergone the teleost-specific 3R WGD and have the fourfold paralogy of a 2R vertebrate (Koh et al., 2003; Noonan et al., 2004; Amemiya et al., 2010). They also allow observed differences between tetrapods and teleosts to be timed specifically to the tetrapod or sarcopterygian lineages (**Figure 1**). Many cryptic vertebrate paralogues have been found in such lineages, and indeed more than fifty developmental genes were found in the coelacanth that had been lost in tetrapods (Amemiya et al., 2013), including *HoxA14* (Amemiya et al., 2010). While these gene losses corresponded to the transition to land, several thousand CNEs near genes involved in morphogenesis and development were identified that evolved anew in the tetrapod lineage (Amemiya et al., 2013).

Looking more closely at developmental genes to address the morphological evolution of limbs in the transition to land, the genomes of coelacanth and lungfish reveal changes to the Hox clusters that may underpin these morphological changes. The coelacanth genome study identified a limb development enhancer in the gene desert upstream of the *HoxD* cluster shared between tetrapods and the coelacanth but absent in the teleost comparison, suggesting this sequence could have been co-opted into lower limb development in tetrapods from an element present in the aquatic ancestral sarcopterygian (Amemiya et al., 2013). In lungfish, *HoxA14* is also present, indicating its loss occurred in tetrapods, and *HoxC13* expression was detected in distal fins, showing the early evidence of its co-option to tetrapod limb development (Meyer et al., 2021). Also implicated in the adaptation to land is an enhancer near *HoxA14* in coelacanth that controls the expression of Posterior Hox genes, though *HoxA14* itself is lost in tetrapods. This enhancer has evolved to be essential for the formation of the placenta (Amemiya et al., 2013). This analysis revealed the impact of new regulatory elements in major transitions in the vertebrates, which are key players in the differentiation among paralogous loci with the potential to evolve different or new functions because of genetic redundancy.

Gene duplications also characterise this transition, particularly in gene families involved in adaptation to land. The lungfish genome contains a gene encoding a respiratory surfactant, *Sftpc*, which originated via duplication earlier in the sarcopterygian lineage, and further duplicated in lungfish and tetrapods (Wang et al., 2021). Also expanded in the lungfish genome are several gene families for olfactory receptors to detect airborne chemicals as well as the vomeronasal receptor family, showing the timing of this adaptation coincides with the adaptation to land (Meyer et al., 2021). Likewise, in comparison with the coelacanth, many tetrapod-specific CNEs were identified near genes involved in olfaction (Amemiya et al., 2013), showing that for the adaptation to land, both gene duplications and changes to gene regulation likely contributed to some extent. These candidate genes and potential regulatory elements may represent some of the genomic changes underpinning the evolution of land vertebrates and were only found with new comparative genomics approaches and new genome sequences. In particular, the lungfish genome is the largest of any animal, so only with new sequencing techniques, in this case ultra-long-read Nanopore technology, could a genome with this quantity of repetitive elements be sequenced (Meyer et al., 2021), enabling a better understanding of the origin of tetrapods.

The Origin of Mammals

A major transition within the tetrapods was the evolution of mammals, most of which are warm-blooded, viviparous, milk-producing, and have fur, distinct from the reptilian ancestor. Of particular relevance to this transition are the monotremes, the basal lineage of mammals who simultaneously exhibit both mammalian and reptilian traits, including milk production and fur, but also have oviparity and produce venom like reptiles, thus providing an excellent point of comparison to address the genetics supporting mammalian adaptations (Warren et al., 2008). There are two extant monotreme lineages, the platypuses and echidnas (**Figure 1**). When the draft genome of the platypus was sequenced, platypuses were found to have many microchromosomes like in reptiles (Warren et al., 2008). Nearly three quarters of the first genome sequence was unmapped to chromosomes, but now, with advancements like long-read sequencing and Hi-C, the platypus genome is assembled to the chromosomal level and the echidna genome was much improved as well (Zhou et al., 2021). These new assemblies allow us to address monotreme traits like their many sex chromosomes, electromagnetic sensory organs, and venom. Mammalian venom evolved convergently to reptilian venoms, originating in monotremes by tandem duplication of defensin family genes in glands that were originally sweat glands, similarly to the evolution of shrew venom and in contrast to snake venom, which originated by duplication of other defensins in salivary glands (Whittington et al., 2008). The many sex chromosomes in monotremes evolved from the same autosomes as bird sex chromosomes (ZW female and ZZ male) rather than the *SRY* (XX female, XY male) sex determination genes of therian mammals, which remain autosomal in monotreme genomes (Veyrunes et al., 2008). The new assembly recharacterised the 10 platypus sex chromosomes and found female-biassed expression of sexually differentiated regions of these chromosomes, consistent with a heterozygous female, and allowed the reconstruction of the series of recombinations and fusions that occurred in their formation (Zhou et al., 2021). This showed that the therian sex determination system evolved independently from the more reptile-like system in monotremes.

A major advancement within mammals was the evolution of viviparity. Monotremes lay small eggs that hatch early in development, as young rely primarily on milk for nutrition, as in therians. The first platypus genome enabled the identification of the genes involved in the transition from oviparity to viviparity, including the retention of reptile-like ZPAX genes and nothepsin, but paired with the loss of vitellogenin genes, leaving only one in the platypus and none in therians, compared to three in birds (Warren et al., 2008). Casein genes are integral to milk production in mammals and evolved via tandem gene duplications of genes in the secretory calcium-binding phosphoprotein family (Kawasaki et al., 2011). Monotremes not only have the full complement of therian caseins, but have a second CSN2 paralogue (Lefèvre et al., 2009). The new assemblies of platypus and echidna genomes have built on these previous findings, including the identification of a CSN3 paralogue and determining the origin of these genes from teeth-related genes based on their syntenic locations (Zhou et al., 2021). A comparative genomics study of the new platypus assembly also revealed that genes involved in the development of the therian placenta were co-opted from genes with other functions, and this transition was underpinned by changes to gene regulation and expression in old genes (Hao et al., 2020). Taken together, these studies show that only with new, chromosome-level assemblies of phylogenetic outgroups can the genetic basis of transitions in chordate evolution be studied.

CONCLUSION

The age of genomics has made sequencing easier and cheaper than ever before, and important advancements such as long-molecule technologies and Hi-C are enabling chromosome-level assemblies for not just traditional model organisms, but also those of particular evolutionary importance. These advances have allowed us to more accurately understand transitions in evolution, for instance using the chromosome-level assembly of *B. floridae* to determine the specific chromosomal rearrangements in and around the 2R WGD. Comparative genomics is now equipped with the material with which to determine the genetic changes that underpin these major evolutionary transitions. It is clear that both gene and genome duplications have been instrumental in the evolution of chordate lineages, from the revolutionary 2R WGD around the origin of the vertebrates, to smaller scale expansions of developmental gene clusters, and tandem duplications underpinning changes to particular lineages or in specific gene families. These duplications can be linked to functional changes as studies of gene expression have revealed the role of regulatory regions, non-coding RNAs, and epigenetics in the evolution of chordate genomes. The continuously improving sequence resources for key outgroup lineages, including amphioxus, cyclostomes, coelacanth and lungfish, and the monotremes, have enabled us to detect gene and genome duplications, pinpoint their timing in the tree of chordates, and to infer the evolutionary impact of these events.

AUTHOR CONTRIBUTIONS

Both authors wrote the manuscript, are accountable for the content, and approved the submitted version.

FUNDING

MEA-R is supported by funding from the University of St Andrews School of Biology and St Leonard's College. Work in the lab of DEKF is funded by the BBSRC (BB/S016856/1) with additional support from the European project Assemble Plus (H2020-INFRAIA-1-2016–2017; grant no. 730984).

ACKNOWLEDGMENTS

We acknowledge discussions with members of the Ferrier laboratory that helped to shape the content of this manuscript.

REFERENCES

Aase-Remedios, M. E., Coll-Lladó, C., and Ferrier, D. E. K. (2020). More Than One-to-Four via 2R: Evidence of an Independent Amphioxus Expansion and Two-Gene Ancestral Vertebrate State for MyoD-Related Myogenic Regulatory Factors (MRFs). *Mol. Biol. Evol.* 37, 2966–2982. doi: 10.1093/molbev/msaa147

Amemiya, C. T., Alfoldi, J., Lee, A. P., Fan, S., Philippe, H., MacCallum, I., et al. (2013). The African coelacanth genome provides insights into tetrapod evolution. *Nature* 496, 311–316. doi: 10.1038/nature12027

Amemiya, C. T., Powers, T. P., Prohaska, S. J., Grimwood, J., Schmutz, J., Dickson, M., et al. (2010). Complete HOX cluster characterization of the coelacanth provides further evidence for slow evolution of its genome. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3622–3627. doi: 10.1073/pnas.0914312107

Amores, A. (1998). Zebrafish hox Clusters and Vertebrate Genome Evolution. *Science* 282, 1711–1714. doi: 10.1126/science.282.5394.1711

Barton-Owen, T. B., Ferrier, D. E. K., and Somorjai, I. M. L. (2018). Pax3/7 duplicated and diverged independently in amphioxus, the basal chordate lineage. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-27700-x

Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., et al. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* 5:3657. doi: 10.1038/ncomms4657

Bertrand, S., Camasses, A., Somorjai, I., Belgacem, M. R., Chabrol, O., Escande, M.-L., et al. (2011). Amphioxus FGF signaling predicts the acquisition of vertebrate morphological traits. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9160–9165. doi: 10.1073/pnas.1014235108

Bertrand, S., and Escriva, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Bliznina, A., Masunaga, A., Mansfield, M. J., Tan, Y., Liu, A. W., West, C., et al. (2021). Telomere-to-telomere assembly of the genome of an individual

Oikopleura dioica from Okinawa using Nanopore-based sequencing. *BMC Genomics* 22:222. doi: 10.1186/s12864-021-07512-6

Braasch, I., Bobe, J., Guiguen, Y., and Postlethwait, J. H. (2018). Reply to: "Subfunctionalization versus neofunctionalization after whole-genome duplication.". *Nat. Genet.* 50, 910–911. doi: 10.1038/s41588-018-0163-3

Braasch, I., Gehrke, A. R., Smith, J. J., Kawasaki, K., Manousaki, T., Pasquier, J., et al. (2016). The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat. Genet.* 48, 427–437. doi: 10.1038/ng.3526

Braun, T., Bober, E., Winter, B., Rosenthal, N., and Arnold, H. H. (1990). Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. *EMBO J.* 9, 821–831.

Brozovic, M., Dantec, C., Dardaillon, J., Dauga, D., Faure, E., Gineste, M., et al. (2018). ANISEED 2017: Extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. *Nucleic Acids Res.* 46, D718–D725. doi: 10.1093/nar/gkx1108

Brunet, F., Crollius, H. R., Paris, M., Aury, J. M., Gibert, P., Jaillon, O., et al. (2006). Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol. Biol. Evol.* 23, 1808–1816. doi: 10.1093/molbev/msl049

Burke, A. C., Nelson, C. E., Morgan, B. A., and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* 121, 333–346.

Butts, T., Holland, P. W. H., and Ferrier, D. E. K. (2008). The Urbilaterian Super-Hox cluster. *Trends Genet.* 24, 259–262. doi: 10.1016/j.tig.2007.09.006

Cañestro, C., Bassham, S., and Postlethwait, J. H. (2003). Seeing chordate evolution through the Ciona genome sequence. *Genome Biol.* 4:208. doi: 10.1186/gb-2003-4-3-208

Cañestro, C., Albalat, R., Irimia, M., and Garcia-Fernàndez, J. (2013). Impact of gene gains, losses and duplication modes on the origin and diversification of vertebrates. *Semin. Cell Dev. Biol.* 24, 83–94. doi: 10.1016/j.semcdb.2012.12.008

Caputo Barucchi, V., Giovannotti, M., Nisi Cerioni, P., and Splendiani, A. (2013). Genome duplication in early vertebrates: Insights from agnathan cytogenetics. *Cytogenet. Genome Res.* 141, 80–89. doi: 10.1159/000354098

Carvajal, J. J., Keith, A., and Rigby, P. W. J. (2008). Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. *Genes Dev.* 22, 265–276. doi: 10.1101/gad.442408

Casaca, A., Santos, A. C., and Mallo, M. (2014). Controlling Hox gene expression and activity to build the vertebrate axial skeleton. *Dev. Dyn.* 243, 24–36. doi: 10.1002/dvdy.24007

Chain, F. J. J., and Evans, B. J. (2006). Multiple Mechanisms Promote the Retained Expression of Gene Duplicates in the Tetraploid Frog Xenopus laevis. *PLoS Genet.* 2:e56. doi: 10.1371/journal.pgen.0020056

Charbonnier, F., Gaspera, B., Della, Armand, A.-S., Van der Laarse, W. J., Launay, T., et al. (2002). Two myogenin-related genes are differentially expressed in Xenopus laevis myogenesis and differ in their ability to transactivate muscle structural genes. *J. Biol. Chem.* 277, 1139–1147. doi: 10.1074/jbc.M107018200

Cheng, P., Huang, Y., Du, H., Li, C., Lv, Y., Ruan, R., et al. (2019). Draft Genome and Complete Hox-Cluster Characterization of the Sterlet (Acipenser ruthenus). *Front. Genet.* 10:776. doi: 10.3389/fgene.2019.00776

Dardaillon, J., Dauga, D., Simion, P., Faure, E., Onuma, T. A., DeBiasse, M. B., et al. (2019). ANISEED 2019: 4D exploration of genetic data for an extended range of tunicates. *Nucleic Acids Res.* 48, D668–D675. doi: 10.1093/nar/gkz955

Dehal, P., and Boore, J. L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLoS Biol.* 3:e314. doi: 10.1371/journal.pbio.0030314

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. doi: 10.1038/nature04336

Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J. M., Da Silva, C., Brinkmann, H., et al. (2010). Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science* 330, 1381–1385. doi: 10.1126/science.1194167

Denton, J. F., Lugo-Martinez, J., Tucker, A. E., Schrider, D. R., Warren, W. C., and Hahn, M. W. (2014). Extensive Error in the Number of Genes Inferred from Draft Genome Assemblies. *PLoS Comput. Biol.* 10:e1003998. doi: 10.1371/journal.pcbi.1003998

Donoghue, P. C. J., and Purnell, M. (2005). Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* 20, 312–319. doi: 10.1016/j.tree.2005.04.008

Du, K., Stöck, M., Kneitz, S., Klopp, C., Woltering, J. M., Adolfi, M. C., et al. (2020). The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization. *Nat. Ecol. Evol.* 4, 841–852. doi: 10.1038/s41559-020-1166-x

Duboule, D. (1994). Temporal colinearity and the phylotypic progression: A basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development* 120, 135–142.

Duboule, D. (2007). The rise and fall of Hox gene clusters. *Development* 134, 2549–2560. doi: 10.1242/dev.001065

Duboule, D., and Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* 8, 1497–1505. doi: 10.1002/j.1460-2075.1989.tb03534.x

Espinosa-Cantú, A., Ascencio, D., Barona-Gómez, F., and DeLuna, A. (2015). Gene duplication and the evolution of moonlighting proteins. *Front. Genet.* 6:227. doi: 10.3389/fgene.2015.00227

Evans, B. J. (2008). Genome evolution and speciation genetics of clawed frogs (Xenopus and Silurana). *Front. Biosci.* 13:4687–4706. doi: 10.2741/3033

Farrè, D., and Albà, M. M. (2010). Heterogeneous patterns of gene-expression diversification in mammalian gene duplicates. *Mol. Biol. Evol.* 27, 325–335. doi: 10.1093/molbev/msp242

Feiner, N., Motone, F., Meyer, A., and Kuraku, S. (2019). Asymmetric paralog evolution between the "cryptic" gene Bmp16 and its well-studied sister genes Bmp2 and Bmp4. *Sci. Rep.* 9:3136. doi: 10.1038/s41598-019-40055-1

Ferrier, D. E. K. (2016). Evolution of Homeobox Gene Clusters in Animals: The Giga-Cluster and Primary vs. Secondary Clustering. *Front. Ecol. Evol.* 4, 1–13. doi: 10.3389/fevo.2016.00036

Ferrier, D. E. K. (2019). "Space and time in Hox / ParaHox gene cluster evolution," in *Perspectives on Evolutionary and Developmental Biology: Essays for Alessandro Minelli*, ed. G. Fusco (Paduva: Padova University Press), 245–258.

Fontana, F., Congiu, L., Mudrak, V. A., Quattro, J. M., Smith, T. I. J., Ware, K., et al. (2008). Evidence of hexaploid karyotype in shortnose sturgeon. *Genome* 51, 113–119. doi: 10.1139/G07-112

Force, A., Amores, A., and Postlethwait, J. H. (2002). Hox cluster organization in the jawless vertebrate Petromyzon marinus. *J. Exp. Zool.* 294, 30–46. doi: 10.1002/jez.10091

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y., and Postlethwait, J. (1999). Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics* 151, 1531–1545.

Fried, C., Prohaska, S. J., and Stadler, P. F. (2003). Independent Hox-cluster duplications in lampreys. *J. Exp. Zool.* 299B, 18–25. doi: 10.1002/jez.b.37

Furlong, R. F., Younger, R., Kasahara, M., Reinhardt, R., Thorndyke, M., and Holland, P. W. H. (2007). A Degenerate ParaHox Gene Cluster in a Degenerate Vertebrate. *Mol. Biol. Evol.* 24, 2681–2686. doi: 10.1093/molbev/msm194

Gans, C., and Northcutt, R. G. (1983). Neural Crest and the Origin of Vertebrates: A New Head. *Science* 220, 268–273. doi: 10.1126/science.220.4594.268

Garcia-Fernàndez, J., and Holland, P. W. H. (1994). Archetypal organization of the amphioxus Hox gene cluster. *Nature* 370, 563–566. doi: 10.1038/370563a0

Gaunt, S. J. (2018). Hox cluster genes and collinearities throughout the tree of animal life. *Int. J. Dev. Biol.* 62, 673–683. doi: 10.1387/ijdb.180162sg

Geraldo, M. T., Valente, G. T., Braz, A. S. K., and Martins, C. (2013). The discovery of Foxl2 paralogs in chondrichthyan, coelacanth and tetrapod genomes reveals an ancient duplication in vertebrates. *Heredity.* 111, 57–65. doi: 10.1038/hdy.2013.19

Glasauer, S. M. K., and Neuhauss, S. C. F. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol. Genet. Genomics* 289, 1045–1060. doi: 10.1007/s00438-014-0889-2

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and Drosophila homeobox gene complexes have common features of organization and expression. *Cell* 57, 367–378. doi: 10.1016/0092-8674(89)90912-4

Green, S. A., and Bronner, M. E. (2014). The lamprey: A jawless vertebrate model system for examining origin of the neural crest and other vertebrate traits. *Differentiation* 87, 44–51. doi: 10.1016/j.diff.2014.02.001

Hao, Y., Lee, H. J., Baraboo, M., Burch, K., Maurer, T., Somarelli, J. A., et al. (2020). Baby Genomics: Tracing the Evolutionary Changes That Gave Rise to Placentation. *Genome Biol. Evol.* 12, 35–47. doi: 10.1093/gbe/evaa026

Hellsten, U., Khokha, M. K., Grammer, T. C., Harland, R. M., Richardson, P., and Rokhsar, D. S. (2007). Accelerated gene evolution and subfunctionalization in the pseudotetraploid frog Xenopus laevis. *BMC Biol.* 5:31. doi: 10.1186/1741-7007-5-31

Hernández-Hernández, J. M., García-González, E. G., Brun, C. E., and Rudnicki, M. A. (2017). The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Semin. Cell Dev. Biol.* 72, 10–18. doi: 10.1016/j.semcdb.2017.11.010

Hoegg, S., Brinkmann, H., Taylor, J. S., and Meyer, A. (2004). Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J. Mol. Evol.* 59, 190–203. doi: 10.1007/s00239-004-2613-z

Hoegg, S., and Meyer, A. (2005). Hox clusters as models for vertebrate genome evolution. *Trends Genet.* 21, 421–424. doi: 10.1016/j.tig.2005.06.004

Holland, L. Z. (2016). Tunicates. *Curr. Biol.* 26, R146–R152. doi: 10.1016/j.cub.2015.12.024

Holland, L. Z., Albalat, R., Azumi, K., Benito-Gutiérrez, È, Blow, M. J., Bronner-Fraser, M., et al. (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* 18, 1100–1111. doi: 10.1101/gr.073676.107

Holland, L. Z., and Holland, N. D. (2001). Evolution of neural crest and placodes: amphioxus as a model for the ancestral vertebrate? *J. Anat.* 199, 85–98. doi: 10.1046/j.1469-7580.199.parts1-2.8.x

Holland, L. Z., and Ocampo Daza, D. (2018). A new look at an old question: when did the second whole genome duplication occur in vertebrate evolution? *Genome Biol.* 19:209. doi: 10.1186/s13059-018-1592-0

Holland, P. W. H. (2013). Evolution of homeobox genes. *Wiley Interdiscip. Rev. Dev. Biol.* 2, 31–45. doi: 10.1002/wdev.78

Hughes, T., Ekman, D., Ardawatia, H., Elofsson, A., and Liberles, D. A. (2007). Evaluating dosage compensation as a cause of duplicate gene retention in Paramecium tetraurelia. *Genome Biol.* 2007:213. doi: 10.1186/gb-2007-8-5-213

Iimura, T., and Pourquié, O. (2006). Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingression. *Nature* 442, 568–571. doi: 10.1038/nature04838

Janvier, P. (2007). Living Primitive Fishes and Fishes From Deep Time. *Fish Physiol.* 26, 1–51. doi: 10.1016/S1546-5098(07)26001-7

Jeffery, W. R., Strickler, A. G., and Yamamoto, Y. (2004). Migratory neural crest-like cells form body pigmentation in a urochordate embryo. *Nature* 431, 696–699. doi: 10.1038/nature02975

Jeffery, W. R. (2006). Ascidian neural crest-like cells: Phylogenetic distribution, relationship to larval complexity, and pigment cell fate. *J. Exp. Zool. Part B Mol. Dev. Evol.* 306, 470–480. doi: 10.1002/jez.b.21109

Jimenez-Delgado, S., Pascual-Anaya, J., and Garcia-Fernandez, J. (2009). Implications of duplicated cis-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty. *Briefings Funct. Genomics Proteomics* 8, 266–275. doi: 10.1093/bfgp/elp029

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomès, D., Rocancourt, D., Buckingham, M., Shinin, V., et al. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* 431, 466–471. doi: 10.1038/nature02876

Kawasaki, K., Lafont, A. G., and Sire, J. Y. (2011). The evolution of milk casein genes from tooth genes before the origin of mammals. *Mol. Biol. Evol.* 28, 2053–2061. doi: 10.1093/molbev/msr020

Kleinjan, D. A., Bancewicz, R. M., Gautier, P., Dahm, R., Schonthaler, H. B., Damante, G., et al. (2008). Subfunctionalization of Duplicated Zebrafish pax6 Genes by cis-Regulatory Divergence. *PLoS Genet.* 4:29. doi: 10.1371/journal.pgen.0040029

Kmita, M., and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301, 331–333. doi: 10.1126/science.1085753

Kobel, H. R., and Du Pasquier, L. (1986). Genetics of polyploid Xenopus. *Trends Genet.* 2, 310–315. doi: 10.1016/0168-9525(86)90286-6

Koh, E. G. L., Lam, K., Christoffels, A., Erdmann, M. V., Brenner, S., and Venkatesh, B. (2003). Hox gene clusters in the Indonesian coelacanth, Latimeria menadoensis. *Proc. Natl. Acad. Sci.* 100, 1084–1088. doi: 10.1073/pnas.0237317100

Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* 78, 191–201. doi: 10.1016/0092-8674(94)90290-9

Krumlauf, R. (2018). Hox genes, clusters and collinearity. *Int. J. Dev. Biol.* 62, 659–663. doi: 10.1387/ijdb.180330rr

Kugler, J. E., Kerner, P., Bouquet, J. M., Jiang, D., and Di Gregorio, A. (2011). Evolutionary changes in the notochord genetic toolkit: A comparative analysis of notochord genes in the ascidian Ciona and the larvacean Oikopleura. *BMC Evol. Biol.* 11:21. doi: 10.1186/1471-2148-11-21

Kuraku, S., Feiner, N., Keeley, S. D., and Hara, Y. (2016). Incorporating tree-thinking and evolutionary time scale into developmental biology. *Dev. Growth Differ.* 58, 131–142. doi: 10.1111/dgd.12258

Kuraku, S., Meyer, A., and Kuratani, S. (2009). Timing of genome duplications relative to the origin of the vertebrates: Did cyclostomes diverge before or after? *Mol. Biol. Evol.* 26, 47–59. doi: 10.1093/molbev/msn222

Ledje, C., Kim, C.-B., and Ruddle, F. H. (2002). Characterization of Hox genes in the Bichir, Polypterus palmas. *J. Exp. Zool.* 294, 107–111. doi: 10.1002/jez.10152

Lefèvre, C. M., Sharp, J. A., and Nicholas, K. R. (2009). Characterisation of monotreme caseins reveals lineage-specific expansion of an ancestral casein locus in mammals. *Reprod. Fertil. Dev.* 21, 1015–1027. doi: 10.1071/RD09083

Lemaire, P. (2011). Evolutionary crossroads in developmental biology: The tunicates. *Development* 138, 2143–2152. doi: 10.1242/dev.048975

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., et al. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4325–4333. doi: 10.1073/pnas.1720115115

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., et al. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature* 533, 200–205. doi: 10.1038/nature17164

Lynch, M., and Conery, J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science* 290, 1151–1155. doi: 10.1126/science.290.5494.1151

Macqueen, D. J., and Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc. R. Soc. B Biol. Sci.* 281:2014. doi: 10.1098/rspb.2013.2881

Marlétaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564, 64–70. doi: 10.1038/s41586-018-0734-6

Martin, K. J., and Holland, P. W. H. (2014). Enigmatic Orthology Relationships between Hox Clusters of the African Butterfly Fish and Other Teleosts Following Ancient Whole-Genome Duplication. *Mol. Biol. Evol.* 31, 2592–2611. doi: 10.1093/molbev/msu202

Martin, K. J., and Holland, P. W. H. (2017). Diversification of Hox gene clusters in osteoglossomorph fish in comparison to other teleosts and the spotted gar outgroup. *J. Exp. Zool. Part B Mol. Dev. Evol.* 328, 638–644. doi: 10.1002/jez.b.22726

Mehta, T. K., Ravi, V., Yamasaki, S., Lee, A. P., Lian, M. M., Tay, B.-H., et al. (2013). Evidence for at least six Hox clusters in the Japanese lamprey (Lethenteron japonicum). *Proc. Natl. Acad. Sci.* 110, 16044–16049. doi: 10.1073/pnas.1315760110

Meyer, A., and Schartl, M. (1999). Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr. Opin. Cell Biol.* 11, 699–704. doi: 10.1016/S0955-0674(99)00039-3

Meyer, A., Schloissnig, S., Franchini, P., Du, K., Woltering, J. M., Irisarri, I., et al. (2021). Giant lungfish genome elucidates the conquest of land by vertebrates. *Nature* 590, 284–289. doi: 10.1038/s41586-021-03198-8

Meyer, A., and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* 27, 937–945. doi: 10.1002/bies.20293

Minguillón, C., Ferrier, D. E. K., Cebrián, C., and Garcia-Fernàndez, J. (2002). Gene duplications in the prototypical cephalochordate amphioxus. *Gene* 287, 121–128. doi: 10.1016/S0378-1119(01)00828-9

Moncaut, N., Rigby, P. W. J., and Carvajal, J. J. (2013). Dial M(RF) for myogenesis. *FEBS J.* 280, 3980–3990. doi: 10.1111/febs.12379

Monteiro, A. S., and Ferrier, D. E. K. (2006). Hox genes are not always Colinear. *Int. J. Biol. Sci.* 2, 95–103. doi: 10.7150/ijbs.2.95

Mungpakdee, S., Seo, H.-C., Angotzi, A. R., Dong, X., Akalin, A., and Chourrout, D. (2008). Differential Evolution of the 13 Atlantic Salmon Hox Clusters. *Mol. Biol. Evol.* 25, 1333–1343. doi: 10.1093/molbev/msn097

Nakai, Y., Kubota, S., Goto, Y., Ishibashi, T., Davison, W., and Kohno, S. (1995). Chromosome elimination in three Baltic, south Pacific and north-east Pacific hagfish species. *Chromosom. Res.* 3, 321–330. doi: 10.1007/BF00713071

Nakai, Y., Kubota, S., and Kohno, S. (1991). Chromatin diminution and chromosome elimination in four japanese hagfish species. *Cytogenet. Genome Res.* 56, 196–198. doi: 10.1159/000133087

Nakatani, Y., Takeda, H., Kohara, Y., and Morishita, S. (2007). Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* 17, 1254–1265. doi: 10.1101/gr.6316407

Navratilova, P., Fredman, D., Hawkins, T. A., Turner, K., Lenhard, B., and Becker, T. S. (2009). Systematic human/zebrafish comparative identification of cis-regulatory activity around vertebrate developmental transcription factor genes. *Dev. Biol.* 327, 526–540. doi: 10.1016/j.ydbio.2008.10.044

Near, T. J., Eytan, R. I, Dornburg, A., Kuhn, K. L., Moore, J. A., Davis, M. P., et al. (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. U. S. A* 109, 13698–13703. doi: 10.1073/pnas.1206625109

Noonan, J. P., Grimwood, J., Danke, J., Schmutz, J., Dickson, M., Amemiya, C. T., et al. (2004). Coelacanth genome sequence reveals the evolutionary history of vertebrate genes. *Genome Res.* 14, 2397–2405. doi: 10.1101/gr.2972804

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin: Springer, doi: 10.1007/978-3-642-86659-3

Pascual-Anaya, J., Sato, I., Sugahara, F., Higuchi, S., Paps, J., Ren, Y., et al. (2018). Hagfish and lamprey Hox genes reveal conservation of temporal colinearity in vertebrates. *Nat. Ecol. Evol.* 2, 859–866. doi: 10.1038/s41559-018-0526-2

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Qian, W., Liao, B. Y., Chang, A. Y. F., and Zhang, J. (2010). Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* 26, 425–430. doi: 10.1016/j.tig.2010.07.002

Robertson, F. M., Gundappa, M. K., Grammes, F., Hvidsten, T. R., Redmond, A. K., Lien, S., et al. (2017). Lineage-specific rediploidization is a mechanism to explain time-lags between genome duplication and evolutionary diversification. *Genome Biol.* 18:111. doi: 10.1186/s13059-017-1241-z

Romanenko, S. A., Biltueva, L. S., Serdyukova, N. A., Kulemzina, A. I., Beklemisheva, V. R., Gladkikh, O. L., et al. (2015). Segmental paleotetraploidy revealed in sterlet (Acipenser ruthenus) genome by chromosome painting. *Mol. Cytogenet.* 8:90. doi: 10.1186/s13039-015-0194-8

Sacerdot, C., Louis, A., Bon, C., Berthelot, C., and Roest Crollius, H. (2018). Chromosome evolution at the origin of the ancestral vertebrate genome. *Genome Biol.* 19, 1–15. doi: 10.1186/s13059-018-1559-1

Santini, F., Harmon, L. J., Carnevale, G., and Alfaro, M. E. (2009). Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol. Biol.* 9:194. doi: 10.1186/1471-2148-9-194

Sato, Y., and Nishida, M. (2010). Teleost fish with specific genome duplication as unique models of vertebrate evolution. *Environ. Biol. Fishes* 88, 169–188. doi: 10.1007/s10641-010-9628-7

Satou, Y., Sato, A., Yasuo, H., Mihirogi, Y., Bishop, J., Fujie, M., et al. (2021). Chromosomal inversion polymorphisms in two sympatric ascidian lineages. *Genome Biol. Evol.* 2021:068. doi: 10.1093/gbe/evab068

Schnapp, E., Pistocchi, A. S., Karampetsou, E., Foglia, E., Lamia, C. L., Cotelli, F., et al. (2009). Induced early expression of mrf4 but not myog rescues myogenesis in the myod/myf5 double-morphant zebrafish embryo. *J. Cell Sci.* 122, 481–488. doi: 10.1242/jcs.038356

Schubert, M., Escriva, H., Xavier-Neto, J., and Laudet, V. (2006). Amphioxus and tunicates as evolutionary model systems. *Trends Ecol. Evol.* 21, 269–277. doi: 10.1016/j.tree.2006.01.009

Schubert, M., Meulemans, D., Bronner-Fraser, M., Holland, L. Z., and Holland, N. D. (2003). Differential mesodermal expression of two amphioxus MyoD family members (AmphiMRF1 and AmphiMRF2). *Gene Expr. Patterns* 3, 199–202. doi: 10.1016/S1567-133X(02)00099-6

Sekigami, Y., Kobayashi, T., Omi, A., Nishitsuji, K., Ikuta, T., Fujiyama, A., et al. (2017). Hox gene cluster of the ascidian, Halocynthia roretzi, reveals multiple ancient steps of cluster disintegration during ascidian evolution. *Zool. Lett.* 3:17. doi: 10.1186/s40851-017-0078-3

Sémon, M., and Wolfe, K. H. (2008). Preferential subfunctionalization of slow-evolving genes after allopolyploidization in Xenopus laevis. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8333–8338. doi: 10.1073/pnas.0708705105

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., et al. (2016). Genome evolution in the allotetraploid frog Xenopus laevis. *Nature* 538, 336–343. doi: 10.1038/nature19840

Simakov, O., Marlétaz, F., Yue, J., O'Connell, B., Jenkins, J., Brandt, A., et al. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nat. Ecol. Evol.* 4, 820–830. doi: 10.1038/s41559-020-1156-z

Smith, J. J., Baker, C., Eichler, E. E., and Amemiya, C. T. (2012). Genetic consequences of programmed genome rearrangement. *Curr. Biol.* 22, 1524–1529. doi: 10.1016/j.cub.2012.06.028

Smith, J. J., and Keinath, M. C. (2015). The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. *Genome Res.* 25, 1081–1090. doi: 10.1101/gr.184135.114

Smith, J. J., Kuraku, S., Holt, C., Sauka-spengler, T., Jiang, N., Campbell, M. S., et al. (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nat. Genet.* 45, 415–421. doi: 10.1038/ng.2568

Smith, J. J., Timoshevskaya, N., Ye, C., Holt, C., Keinath, M. C., Parker, H. J., et al. (2018). The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat. Genet.* 50, 270–277. doi: 10.1038/s41588-017-0036-1

Somorjai, I., Bertrand, S., Camasses, A., Haguenauer, A., and Escriva, H. (2008). Evidence for stasis and not genetic piracy in developmental expression patterns of Branchiostoma lanceolatum and Branchiostoma floridae, two amphioxus species that have evolved independently over the course of 200 Myr. *Dev. Genes Evol.* 218, 703–713. doi: 10.1007/s00427-008-0256-6

Steinke, D., Salzburger, W., Braasch, I., and Meyer, A. (2006). Many genes in fish have species-specific asymmetric rates of molecular evolution. *BMC Genomics* 7, 1–18. doi: 10.1186/1471-2164-7-20

Stolfi, A., Ryan, K., Meinertzhagen, I. A., and Christiaen, L. (2015). Migratory neuronal progenitors arise from the neural plate borders in tunicates. *Nature* 527, 371–374. doi: 10.1038/nature15758

Tan, X., Zhang, P. J., and Du, S. J. (2014). Evolutionary aspects of a new MyoD gene in amphioxus (Branchiostoma belcheri) and its promoter specificity in skeletal and cardiac muscles. *Biol.* 69, 1210–1221. doi: 10.2478/s11756-014-0427-z

Tapscott, S. J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development* 132, 2685–2695. doi: 10.1242/dev.01539

Taylor, J. S., Van De Peer, Y., and Meyer, A. (2001). Genome duplication, divergent resolution and speciation. *Trends Genet.* 17, 299–301. doi: 10.1016/S0168-9525(01)02318-6

Timoshevskiy, V. A., Lampman, R. T., Hess, J. E., Porter, L. L., and Smith, J. J. (2017). Deep ancestry of programmed genome rearrangement in lampreys. *Dev. Biol.* 429, 31–34. doi: 10.1016/j.ydbio.2017.06.032

Tschopp, P., Tarchini, B., Spitz, F., Zakany, J., and Duboule, D. (2009). Uncoupling Time and Space in the Collinear Regulation of Hox Genes. *PLoS Genet.* 5:e1000398. doi: 10.1371/journal.pgen.1000398

Uno, Y., Nishida, C., Takagi, C., Ueno, N., and Matsuda, Y. (2013). Homoeologous chromosomes of Xenopus laevis are highly conserved after whole-genome duplication. *Heredity* 111, 430–436. doi: 10.1038/hdy.2013.65

Urano, A., Suzuki, M. M., Zhang, P., Satoh, N., and Satoh, G. (2003). Expression of muscle-related genes and two MyoD genes during amphioxus notochord development. *Evol. Dev.* 5, 447–458. doi: 10.1046/j.1525-142X.2003.03051.x

Vasil'yev, V. P., Solokov, L. I., and Serebryakova, Y. V. (1980). Karyotypes of the Siberian sturgeon, Acipenser baeri, of the Lena river and some aspects of karyotype evolution in Acipenseriformes. *J. Ichthyol.* 20, 37–45.

Veyrunes, F., Waters, P. D., Miethke, P., Rens, W., McMillan, D., Alsop, A. E., et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res.* 18, 965–973. doi: 10.1101/gr.7101908

Wang, K., Wang, J., Zhu, C., Yang, L., Ren, Y., Ruan, J., et al. (2021). African lungfish genome sheds light on the vertebrate water-to-land transition. *Cell* 184, 1362.e–1376.e. doi: 10.1016/j.cell.2021.01.047

Warren, W. C., Hillier, L. D. W., Marshall Graves, J. A., Birney, E., Ponting, C. P., Grützner, F., et al. (2008). Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453, 175–183. doi: 10.1038/nature06936

Werner, H. B. (2013). Do we have to reconsider the evolutionary emergence of myelin? *Front. Cell. Neurosci.* 6, 415–421. doi: 10.3389/fncel.2013.00217

Whittington, C. M., Papenfuss, A. T., Bansal, P., Torres, A. M., Wong, E. S. W., Deakin, J. E., et al. (2008). Defensins and the convergent evolution of platypus and reptile venom genes. *Genome Res.* 18, 986–994. doi: 10.1101/gr.7149808

Yu, J.-K., Meulemans, D., McKeown, S. J., and Bronner-Fraser, M. (2008). Insights from the amphioxus genome on the origin of vertebrate neural crest. *Genome Res.* 18, 1127–1132. doi: 10.1101/gr.076208.108

Yuan, J., Zhang, S., Liu, Z., Luan, Z., and Hu, G. (2003). Cloning and phylogenetic analysis of an amphioxus myogenic bHLH gene AmphiMDF. *Biochem. Biophys. Res. Commun.* 301, 960–967. doi: 10.1016/S0006-291X(03)00081-0

Zhang, J. (2003). Evolution by gene duplication: An update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/S0169-5347(03)00033-8

Zhang, Q.-L., Zhang, G.-L., Yuan, M.-L., Dong, Z.-X., Li, H.-W., Guo, J., et al. (2018). A Phylogenomic Framework and Divergence History of Cephalochordata Amphioxus. *Front. Physiol.* 9:1833. doi: 10.3389/fphys.2018.01833

Zhou, Y., Shearwin-Whyatt, L., Li, J., Song, Z., Hayakawa, T., Stevens, D., et al. (2021). Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* 2021, 1–7. doi: 10.1038/s41586-020-03039-0

Zhu, T., Li, Y., Pang, Y., Han, Y., Li, J., Wang, Z., et al. (2021). Chromosome-level genome assembly of Lethenteron reissneri provides insights into lamprey evolution. *Mol. Ecol. Resour.* 21, 448–463. doi: 10.1111/1755-0998.13279

Conflict of Interest: 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 © 2021 Aase-Remedios and Ferrier. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 02 July 2021
doi: 10.3389/fcell.2021.702688

Modular Evolution and Population Variability of *Oikopleura dioica* Metallothioneins

Sara Calatayud[1]†, Mario Garcia-Risco[2]†, Mercè Capdevila[2], Cristian Cañestro[1], Òscar Palacios[2]* and Ricard Albalat[1]*

[1] Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain, [2] Departament de Química, Facultat de Ciències, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
María Ayelén Pagani,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
Filipe Castro,
University of Porto, Portugal

***Correspondence:**
Òscar Palacios
Oscar.Palacios@uab.cat
Ricard Albalat
ralbalat@ub.edu

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *29 April 2021*
Accepted: *09 June 2021*
Published: *02 July 2021*

Citation:
Calatayud S, Garcia-Risco M, Capdevila M, Cañestro C, Palacios Ò and Albalat R (2021) Modular Evolution and Population Variability of Oikopleura dioica Metallothioneins. Front. Cell Dev. Biol. 9:702688. doi: 10.3389/fcell.2021.702688

Chordate *Oikopleura dioica* probably is the fastest evolving metazoan reported so far, and thereby, a suitable system in which to explore the limits of evolutionary processes. For this reason, and in order to gain new insights on the evolution of protein modularity, we have investigated the organization, function and evolution of multi-modular metallothionein (MT) proteins in *O. dioica*. MTs are a heterogeneous group of modular proteins defined by their cysteine (C)-rich domains, which confer the capacity of coordinating different transition metal ions. *O. dioica* has two MTs, a bi-modular OdiMT1 consisting of two domains (t-12C and 12C), and a multi-modular OdiMT2 with six t-12C/12C repeats. By means of mass spectrometry and spectroscopy of metal-protein complexes, we have shown that the 12C domain is able to autonomously bind four divalent metal ions, although the t-12C/12C pair –as it is found in OdiMT1– is the optimized unit for divalent metal binding. We have also shown a direct relationship between the number of the t-12C/12C repeats and the metal-binding capacity of the MTs, which means a stepwise mode of functional and structural evolution for OdiMT2. Finally, after analyzing four different *O. dioica* populations worldwide distributed, we have detected several OdiMT2 variants with changes in their number of t-12C/12C domain repeats. This finding reveals that the number of repeats fluctuates between current *O. dioica* populations, which provides a new perspective on the evolution of domain repeat proteins.

Keywords: appendicularian tunicate chordates, modular protein evolution, metallothionein domains, tandem domain repeats, population variants, intra-species variability

INTRODUCTION

Oikopleura dioica is a tunicate species of the appendicularian class in the chordate phylum. This species is emerging as a non-classical animal model in the field of evolutionary developmental biology (a.k.a. evo-devo) especially attractive for its unusually dynamic gene and genome evolution (reviewed in Ferrández-Roldán et al., 2019). At genome level, *O. dioica* has suffered numerous chromosomal rearrangements accompanied by a drastic process of compaction, becoming the smallest known chordate genome (Denoeud et al., 2010). At gene level, besides an extraordinary amount of gene duplications and losses, *O. dioica* sequences show high evolutionary rates, which are on average two-three times higher than in ascidians and vertebrates (Berna et al., 2012; Berna and Alvarez-Valin, 2014). *O. dioica* probably is the fastest evolving metazoan reported so far (Edvardsen et al., 2005; Denoeud et al., 2010). In addition, its pattern of amino acid substitution also shows some unusual traits in comparison with other chordates. Cysteines (Cys, C), for instance, are the less conserved amino acids in *O. dioica* proteins (Berna et al., 2012; Berná and Alvarez-Valin, 2015), whereas Cys are one of the most conserved amino acids according with classical analyses of protein evolution (Henikoff and Henikoff, 1992; Marino and Gladyshev, 2010). These exceptional evolutionary features make comparative studies between *O. dioica* and other chordate species suitable for understanding the functional and structural limits to which chordate genes and proteins can evolve.

In protein evolution, domains are considered evolutionary modules, and the majority of proteins of all living beings are multi-modular proteins that consist of several domains (Apic et al., 2001). While the creation of multi-modular proteins through shuffling of different domain types has been extensively analyzed (Apic et al., 2001; Björklund et al., 2005; Dohmen et al., 2020), the evolution of proteins made of tandem domain repeats is less understood (Björklund et al., 2006). To get new insights into the functional and structural evolution of these domain repeat proteins, we have focused on the multi-modular metallothioneins (MTs), using those of the fast evolving *O. dioica* species as case study.

Metallothioneins are proteins known for their high percentage of cysteines (Kojima et al., 1999), which confers them the capability of binding both essential and non-essential transition metals (reviewed in Capdevila et al., 2012; Blindauer, 2014). The Cys residues of MTs are arranged in distinctive motifs (i.e., CxC, CC, and CCC), whose number and distribution led to define different functional domains, originally for mammalian MTs (Braun et al., 1986), and later, in other animal MTs (Riek et al., 1999; Munoz et al., 2002; Baumann et al., 2017; Beil et al., 2019; Calatayud et al., 2021b). In chordates, for instance, vertebrate and cephalochordate MTs are bi-modular proteins with two domains that have diverse preferences and capacities for binding zinc (Zn), copper (Cu), or cadmium (Cd) ions (Capdevila and Atrian, 2011; Vasak and Meloni, 2011; Guirola et al., 2012; Artells et al., 2013). In contrast, most tunicate MTs are mono-modular proteins, whose single domain has a pervasive preference for Cd(II) ions (Calatayud et al., 2021a). The domain configuration of each MT is, indeed, functionally and structurally relevant because domains determine the formation of metal-thiolate clusters: domains with 9 Cys cluster with three divalent metal ions, while 11/12 Cys domains cluster with four divalent metal ions (e.g., mammalian β and α domains, respectively, Otvos and Armitage, 1980; Schultze et al., 1988). In addition, domain analyses have been shown to be helpful for elucidating the origin and evolutionary relationships of MTs in diverse groups of mollusks (Jenny et al., 2016; Nam and Kim, 2017; Calatayud et al., 2021b), and to reconstruct the complex evolutionary history of chordate MTs (Calatayud et al., 2021a).

O. dioica has two MTs, a bi-modular OdiMT1 and a multi-modular OdiMT2 (formerly OdMT1 and OdMT2) made of different number of domain repeats (Calatayud et al., 2018). The arrangement of Cys motifs in *O. dioica* domains diverges from that found in the MTs of other tunicates belonging to the ascidian and thaliacean classes, but it is similar to that of other appendicularian species of the same genus, *O. albicans* and *O. vanhoeffeni* (Calatayud et al., 2021a). Comparison of the appendicularian MTs show that the original *Oikopleura* MT domain had twelve cysteines (12C), and that this domain corresponds to previously described C7 + C5 subunits (Calatayud et al., 2018, 2021a). OdiMT1 would have therefore two 12C domains, but its amino-terminal domain was "trimmed" to become a t-12C domain that lacks the C5 subunit. OdiMT2 would be a multi-modular MT derived from an ancestral copy with a t-12C/12C domain organization, similar to the current OdiMT1, that was tandem duplicated five times yielding its domain repeat organization (Calatayud et al., 2018). In this work, we have pursued the analysis of *O. dioica* MTs, paying special attention to their modular configuration. First, we have characterized the metal-binding features of the original *Oikopleura* 12C domains –both the full-length (12C) and the trimmed (t-12C)–, revealing that although the 12C domain autonomously coordinates divalent metal ions, the t-12C/12C pair seems an improved form for divalent metal binding. Second, we have shown a direct relationship between the number of the t-12C/12C domain repeats and the metal-binding capacity of OdiMTs. Finally, taking advantage of the high level of genetic variation among *O. dioica* populations (Wang et al., 2020; Bliznina et al., 2021), we have detected population variants of OdiMT2 with changes in their number of t-12C/12C domains. Our data expose a high structural plasticity of MTs in *O. dioica* that, as if it was a natural test-bench, seems to be exploring the chordate limits of MT modularity.

MATERIALS AND METHODS

Production and Purification of Recombinant Metal-MT Complexes

Production and purification of recombinant metal-MT complexes of proteins with different number and combinations of t-12C and 12C domains (see **Table 1** for details; sequences of the domains are from Norwegian OdiMT1 and OdiMT2 sequences) was performed as described elsewhere (Calatayud et al., 2018, 2021b). In brief, synthetic cDNAs codifying the different constructs of 12-Cys domains were provided by Synbio Technologies (Monmouth Junction, NJ, United States), cloned in the pGEX-4T-1 expression vector (GE Healthcare, Chicago, IL, United States) and transformed in protease-deficient *E. coli* BL21 strain. Metal-MT complexes were produced in *E. coli* BL21 cultures expressing the recombinant plasmids, after induction with isopropyl-β-D-thiogalactopyranoside (100 μM) and supplementation with $ZnCl_2$ (300 μM), $CdCl_2$ (300 μM), or $CuSO_4$ (500 μM). Metal-MT complexes were purified from the soluble protein fraction of sonicated bacteria by affinity purification of the GST-tagged proteins, and digestion with thrombin. Notice that the digestion with thrombin added two additional residues, Gly and Ser, at the N-terminal end of all purified proteins. These two amino acids do not interfere with the metal-binding features of recombinant MTs (Cols et al., 1997). The metal-MT complexes were concentrated with a 3 kDa Centripep Low Concentrator (Amicon, Merck), and fractionated on a Superdex-75 FPLC column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.0 or with fresh 50 mM ammonium acetate, pH 7.0, and run at 0.8 mL min^{-1}. The protein-containing fractions, identified by their absorbance at 254 nm, were pooled and stored at −80°C until use.

Analysis of Metal-MT Complexes

All designed constructs of OdiMTs were characterized by means of mass spectrometry (ESI-MS) and spectroscopy (ICP-AES). An electrospray ionization mass spectrometry (ESI-MS) Micro Tof-Q Instrument (Brucker Daltonics Gmbh, Bremen, Germany) interfaced with a Series 1100 HPLC pump (Agilent Technologies) was used to determine the molecular mass of the recombinant proteins. The instrument was calibrated with ESI-L Low Concentration Tuning Mix (Agilent Technologies, United States) and the experimental conditions were set up as follows: injection of 10–20 μL of sample through a PEEK long tube (1–1.5 m × 0.18 mm i.d.) at 30–50 μL min^{-1}; capillary-counterelectrode voltage, 3.5–5.0 kV; desolvation temperature, 90–110°C; dry gas, 6 L min^{-1}. Data was acquired over an *m/z* range of 800–3,000. The liquid carriers were a 90:10 mixture of 15 mM ammonium acetate and acetonitrile at pH 7.0 and a 95:5 mixture of formic acid and acetonitrile at pH 2.4.

Element concentrations of S, Zn, Cd, and Cu in the sample were determined by Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES) by means of a Perkin-Elmer Optima 4300 DV (Waltham, United States) at the correct wavelength (S, 182.04 nm; Zn, 213.86 nm; Cd, 228.80 nm; Cu, 324.80 nm) under conventional conditions (Bongers et al., 1988). MTs concentration was calculated based on the S concentration obtained by ICP-AES, assuming that all the sulfur measured comes from peptides' Cys and Met residues.

Analysis of MT Variation

We analyzed the MT sequence variations of four geographically distant *O. dioica* populations: Norway, Japan (Osaka and Okinawa), Oregon, and Catalonia. We used Norwegian OdiMT1 (NCBI accession number CABV01001936.1) and OdiMT2 (NCBI accession number CABV01001042.1) sequences retrieved from the Oikobase genome database[1] (Denoeud et al., 2010) as reference for blast searches in ANISEED[2] and NCBI Sequence Read Archives[3] for Japanese population (Osaka and Okinawa, respectively). Raw sequence data from SRA searches was assembled using SeqMan 8.0.2 (Pro Assembler) software from the DNASTAR Lasergene package, and manually inspected in order to reconstruct the MT sequences.

For Catalonian sequences, we PCR amplified the *MT* genes from specimens captured in the Mediterranean coast of Barcelona and cultured in our animal facility at the University of Barcelona (Martí-Solans et al., 2015). Primers and PCR conditions are listed in **Supplementary Table 1**. PCR products were cloned using Topo TA Cloning® Kit of Invitrogen and sequenced at the Scientific and Technological Centers of the University of Barcelona, using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) in an automatic sequencer (ABIPRISM 310, Applied Biosystems). For Oregonian sequences, we took advantage of a data of a genomic shotgun approach. All new MT sequences have been deposited in Genbank database at NCBI[4].

For comparisons, MT sequences from different *O. dioica* populations were aligned with Aliview program (Larsson, 2014) and reviewed manually. Phylogenetic reconstructions were based on ML inferences calculated with PhyML v3.0 and automatic mode of selection of substitution model (Guindon et al., 2010) using protein sequence alignments. Sequences and the accession numbers used in this study are provided in **Supplementary Table 2**.

RESULTS AND DISCUSSION

Functional Analysis of the Modularity in OdiMTs

Metal-Binding Capabilities of 12C and t-12C Domains

OdiMT1 is a 72 amino acid long protein (20 Cys, 28%) that can be divided in three parts: an N-terminal peptide made of six amino acids; next, an amino-terminal t-12C domain; and finally, a carboxyl-terminal 12C domain. OdiMT2 is a protein of 399 amino acids (123 Cys, 31%) that consists of the N-terminal peptide, and six tandem repeat units (RU1–RU6), each one made of a t-12C domain and a 12C domain (**Figure 1**; Calatayud et al., 2018). To investigate the functionality of this modular organization, we studied the formation of metal-MT complexes of recombinant proteins containing different combinations of t-12C and 12C domains (**Table 1**). The metal-binding capability of each construct was analyzed by inductively coupled plasma atomic emission spectroscopy (ICP-AES), and the species recovered identified by electrospray ionization mass spectrometry (ESI-MS) (Capdevila et al., 2012). ICP-AES was also used for metal-to-protein stoichiometry determination through the measurement of element composition of the samples (S, Zn, Cd, and Cu) (Bongers et al., 1988).

We first characterized the metal-binding features of the 12C domain by analyzing its ability to form metal complexes when expressed alone. Based on OdiMT1 sequence, we designed construct 1.1, which encoded a 42 amino acid peptide (residues from 31 to 72) comprising the 12C domain of OdiMT1 (**Table 1**). Notice that the 12C domain of OdiMT1 is, indeed, an 11C domain due to a Cys to Ser substitution in the carboxyl terminal region in comparison with prototypical 12C domain (thus, we named this domain as 11C/12C; **Figure 1**). Our results showed that this 11C/12C domain was able to yield almost unique Zn_4- and Cd_4-protein species (**Figure 2**), in a similar way that other domains with 11 or 12 Cys: the 11C mammal and echinoderm

---

[1] http://oikoarrays.biology.uiowa.edu/Oiko

[2] https://www.aniseed.cnrs.fr/

[3] https://www.ncbi.nlm.nih.gov/sra/

[4] https://www.ncbi.nlm.nih.gov/genbank/

TABLE 1 | Heterogously produced proteins containing different number and combinations of t-12C and 12C domains.

Construct	Scheme[a]	Domains	Amino acid included (Num. C)	Num.aa[b]	Major metal-protein species
OdiMT (full protein)[c]		**N-term + t-12C + 12C**	From 1 to 72 (20)	72	$Cd_7/Zn_7/Cu_{12}$
OdiMT1-Construct 1.1		**12C**	From 31 to 72 (11)	42	$Cd_4/Zn_4/Cu_9$
OdiMT1-Construct 1.2		**N-term + 12C**	From 1 to 6 (1) + from 31 to 72 (11)	48	$Cd_5/Zn_2/Cu_4$
OdiMT1-Construct 1.3		**N-term + 12C**	From 1 to 30 (9) + from 55 to 72 (4)	48	$Cd_5/Zn_4/Cu_{11}$
OdiMT1-Construct 1.4		**t-12C**	From 7 to 30 (8)	24	Cd_3/Zn_3/nd[d]
OdiMT1-Construct 1.5		**t-12C**	From 31 to 54 (7)	24	Cd_3/Zn_3/nd
OdiMT2 (full protein)[c]	t12C 12C t12C 12C t12C 12C t12C 12C t12C 12C t12C 12C	**N-term + [t-12C+12C] ×6**	From 1 to 399 (123)	399	nd/nd/nd
OdiMT2-Construct 2.1	t12C 12C t12C 12C t12C 12C	**N-term + [t-12C+12C] ×3**	From 1 to 199 (61) + from 396 to 399 (2)	203	nd/Cd_{22}/nd
OdiMT2-Construct 2.2	t12C 12C t12C 12C t12C 12C	**[t-12C+12C] ×3**	From 201 to 399 (62)	199	Zn_{20}/Cd_{22}/nd

[a] Colored boxes represents the six amino acids of the N-terminus (black), a "trimmed" 12C domain (green) and a complete 12C domain (orange).
[b] This number does not include two additional amino acids (Gly and Ser) added at the N-terminus of the recombinant proteins as consequence of the experimental design for the expression and purification of the metal-protein complexes (see section "Materials and Methods" for details).
[c] OdiMT1 and OdiMT2 full proteins were analyzed in Calatayud et al. (2018).
[d] nd, no detected.

α domains (Stillman et al., 1987; Tomas et al., 2013), the 12C mollusk α domains (Digilio et al., 2009), the insect 12C MTs (Egli et al., 2006) and the ascidian and thaliacean 12C MTs (Calatayud et al., 2021a). These results revealed a significant structural and functional autonomy of *O. dioica* 11C/12C domain, which was able to form stable metal-protein clusters with Zn(II) and Cd(II). Regarding the metal preference of this domain, the mixture of multiple Cu_n-protein complexes (n ranging from 4 to 10) in Cu(II) surplus productions (**Figure 2**) discarded a preference of this domain for monovalent Cu(I) ions. In addition, a preference for Cd(II) over Zn(II) might be indirectly inferred not only from the neatness of the ESI-MS spectra of the Cd-preparation but also from the fact that the domain rendered homometallic Cu(I) species in the Cu-preparations (**Figure 2**), which is characteristic of Cd-thioneins, whereas Zn-thioneins yield heterometallic Zn/Cu-MT complexes when expressed Cu-enriched media (Palacios et al., 2011). Overall, we concluded that the 11C/12C domain of OdiMTs formed stable clusters with four Cd(II) ions, which was in agreement with the biochemical features reported for the full OdiMT1 protein (Calatayud et al., 2018).

Although we obtained reliable results for the 11C/12C domain of OdiMT1, we wondered if the extra Cys found the N-terminal peptide of OdiMT1 could compensated the loss of one Cys in this domain, significantly improving its metal coordination features (i.e., enhancing specificity, increasing stability or metal-to-protein stoichiometries). To test this possibility, we designed two constructs: construct 1.2, which expressed a 48 amino acid peptide comprising the six N-term residues of OdiMT1 (which included the extra Cys) fused to the 11C/12C domain (**Table 1**); and construct 1.3, which also expressed a 48 amino acid peptide comprising the six N-term residues fused to a 12C domain resulting from the combination of the t-12C domain of OdiMT1 (from 7 to 30) with the 18 last residues of the carboxyl-terminus of the protein (from 55 to 72) (**Table 1**). Although both constructs rendered Cd_5-protein complexes as major species according to ESI-MS data (**Figure 2**), the samples recovered from both, Zn- and Cd-supplemented cultures, resulted to be a significant mixture of metal-protein complexes and some Cd-protein species contained sulfide labile ligands (**Figure 2**). The presence of S^{2-} ions is probably due to the incapability of these artificially designed peptides to build a stable metal cluster as these "extra" ligands can aid in the stabilization of the metal clusters (Capdevila et al., 2005). This, together with the heterogeneity of the samples suggested, therefore, that the 1.2 and 1.3 constructs have not improved the metal binding abilities of construct 1.1, questioning the contribution of the extra Cys in the N-terminal peptide to the metal coordination, and reinforcing the functional entity of the 11C/12C domain as an efficient solution for coordinating divalent metal ions emerged during the evolution of the *Oikopleura* lineage.

Next, we investigated the functionality of the t-12C domain, which is a 12C domain lacking the carboxyl C5 subunit, and thereby, containing only 7/8 cysteines. We analyzed the metal-binding features of two different t-12C domains expressed by two constructs. Construct 1.4 produced the t-12C domain of OdiMT1, from residues 7 to 30. Construct 1.5 produced the t-12C domain resulting from the truncation of the last 18 residues of the 11/12C domain of OdiMT1, from residues 31 to 54 (**Table 1**). Our analyses showed that both constructs mainly bind three divalent metal ions, either Zn(II) or Cd(II) (**Figure 2**). This metal-to-protein stoichiometry agreed with the possibility that the seven divalent metal ions (M_7) coordinated by the full OdiMT1 protein are organized in two metal clusters: an 11C/12C domain cluster with four metal ions ($M^{II}{}_4$) at the carboxyl-end of the MT, and a t-12C domain cluster with three metal ions ($M^{II}{}_3$) at the amino-term region. The t-12C domains rendered, however, multiple metallospecies and some Cd-protein species

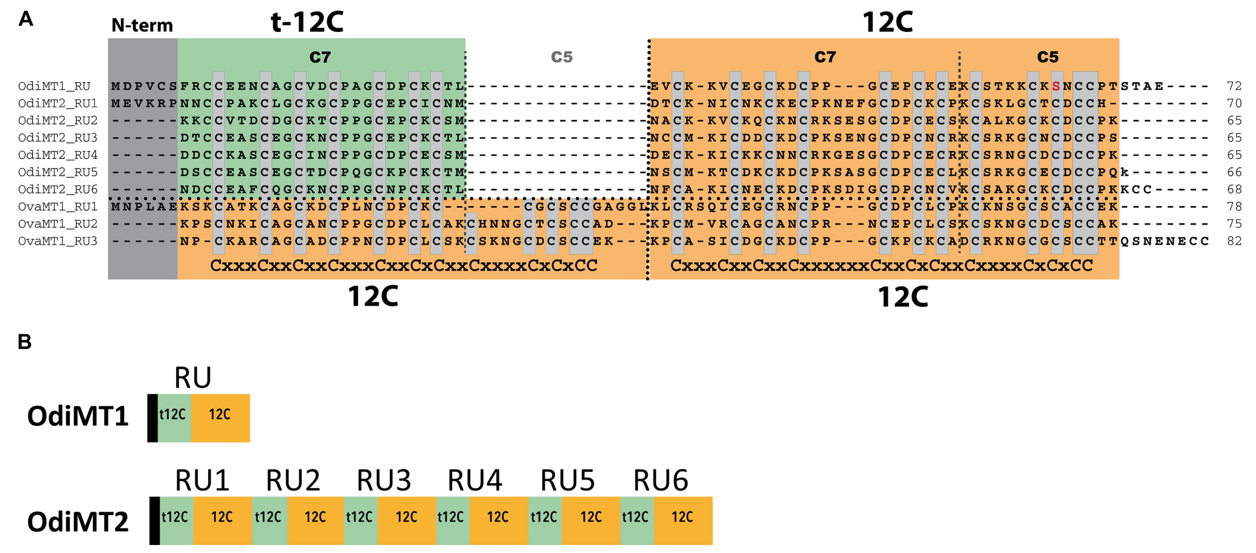


FIGURE 1 | Domain organization of *O. dioica* metallothioneins. **(A)** Amino acid alignment of *O. dioica* OdiMT1 and OdiMT2 split in its repeat units (RU1–RU6). The sequence of *O. vanhoeffeni* MT (OvaMT1) split in its repeat units (RU1–RU3) has been included to show that the amino-terminal 12C domains of OdiMT1 and of the RU of OdiMT2 are "trimmed" (t-12C; green background), and lacks the C5 subunit. Conserved cysteines are highlighted in gray. The arrangement of Cys motifs of each 12C domain is showed below the alignment. Notice that the 12C domain (orange background) corresponds to previously described C7 + C5 subunits (Calatayud et al., 2018). **(B)** Schematic representation of the modular organization of OdiMT1 and OdiMT2 proteins. OdiMT1 consists of 6 amino acids of the N-terminus (black box) followed by a single repeat unit (RU) with a t-12C domain (green box) and a 12C domain (orange box), while OdiMT2 is made of six RU, RU1–RU6.

with sulfide ligands (**Figure 2**), which suggested that the t-12C domain expressed alone did not efficiently bind the metals by itself, and that it would require the neighboring 11C/12C domain to properly coordinate the seven metal ions in the full MT protein (Calatayud et al., 2018).

Metal-Binding Capabilities of t-12C/12C Repeats

We also analyzed the metal-binding features of tandem repeats of t-12C/12C domains as they are found in OdiMT2. We designed two constructs. Construct 2.1, encoding a 203 amino acid protein, comprised the three amino-terminal RU (i.e., RU$_1$RU$_2$RU$_3$; residues 1–199), plus the four last amino acids (two of them Cys) of the carboxyl-end (from 396 to 399). Construct 2.2, encoding a 199 amino acid protein, comprised the three carboxyl-terminal RU (i.e., RU$_4$RU$_5$RU$_6$) plus the four last amino acids (from 201 to 399) (**Figure 1** and **Table 1**). Recombinant synthesis of the two partial OdiMT2 proteins –RU$_1$RU$_2$RU$_3$ and RU$_4$RU$_5$RU$_6$– in Zn(II)- and Cd(II)-enriched *E. coli* cultures rendered a low yield of metal-protein complexes, but ICP-AES analyses showed that both produced complexes with divalent metal ions, mostly with Cd(II) ions. ESI-MS analyses (**Figure 3**) showed mixtures of metal-protein complexes with different stoichiometries, and Cd-protein complexes with sulfide ligands, being Zn$_{20-21}$, Cd$_{22-23}$, and Cd$_{23-24}$S-complexes the major species (**Figure 3**). The metal-to-protein stoichiometry of these species was close to what was expected considering that the partial proteins consist of three RU, each one equivalent to a full-length OdiMT1 that binds seven Zn(II)/Cd(II) ions, i.e., 3 RU × 7 M^{II} = 21 total divalent metal ions. The two additional Cys at the carboxyl-end, the presence of sulfide ligands, and the fact that multiple metallospecies are common in multi-modular MT productions (Palacios et al., 2014; Iturbe-Espinoza et al., 2016) could account for the slightly high metal content (up to 24 ions) observed in some of them. Interestingly, ESI-MS spectra of both constructs at pH 2.4 showed species loaded with 4 and 8 divalent metal ions (**Figure 3**). These species might be reflecting the t-12C/12C organization of each RU, in which the "standard" 12C domain would be more reluctant to release its four metal ions than the "trimmed" t-C12 domain. Overall, the metal stoichiometry of the partial MTs pointed to a direct relationship between the number of the t-12C/12C repeats and the metal-binding capacity of multi-modular OdiMTs.

In conclusion, our analyses suggested that during the evolution of MTs in *O. dioica*, an ancestral 12C domain was tandem duplicated. The N-terminal 12C copy was trimmed (t-12C domain), partially losing its autonomy for metal binding, and the t-12C/12C pair became the optimized functional unit. Afterward, this primeval t-12C/12C *OdiMT* gene was duplicated. While one of the duplicates remained unaltered as the current *OdiMT1*, the other copy suffered several internal tandem duplications of the functional t-12C/12C pair in an evolutionary process that stepwise changed the number of domain repeats affecting the metal binding capacity of the new multi-modular OdiMT2. Domain expansions that generate high metal-binding capacity MTs such as OdiMT2, gene duplications that lead to the amplification of the number of *MT* genes such as those of *CUP1* in yeast (Adamo et al., 2012), and mutations in regulatory regions that lead to the overexpression of MTs in insects (Costa et al., 2012; Catalan et al., 2016) appear to be different ways of increasing the

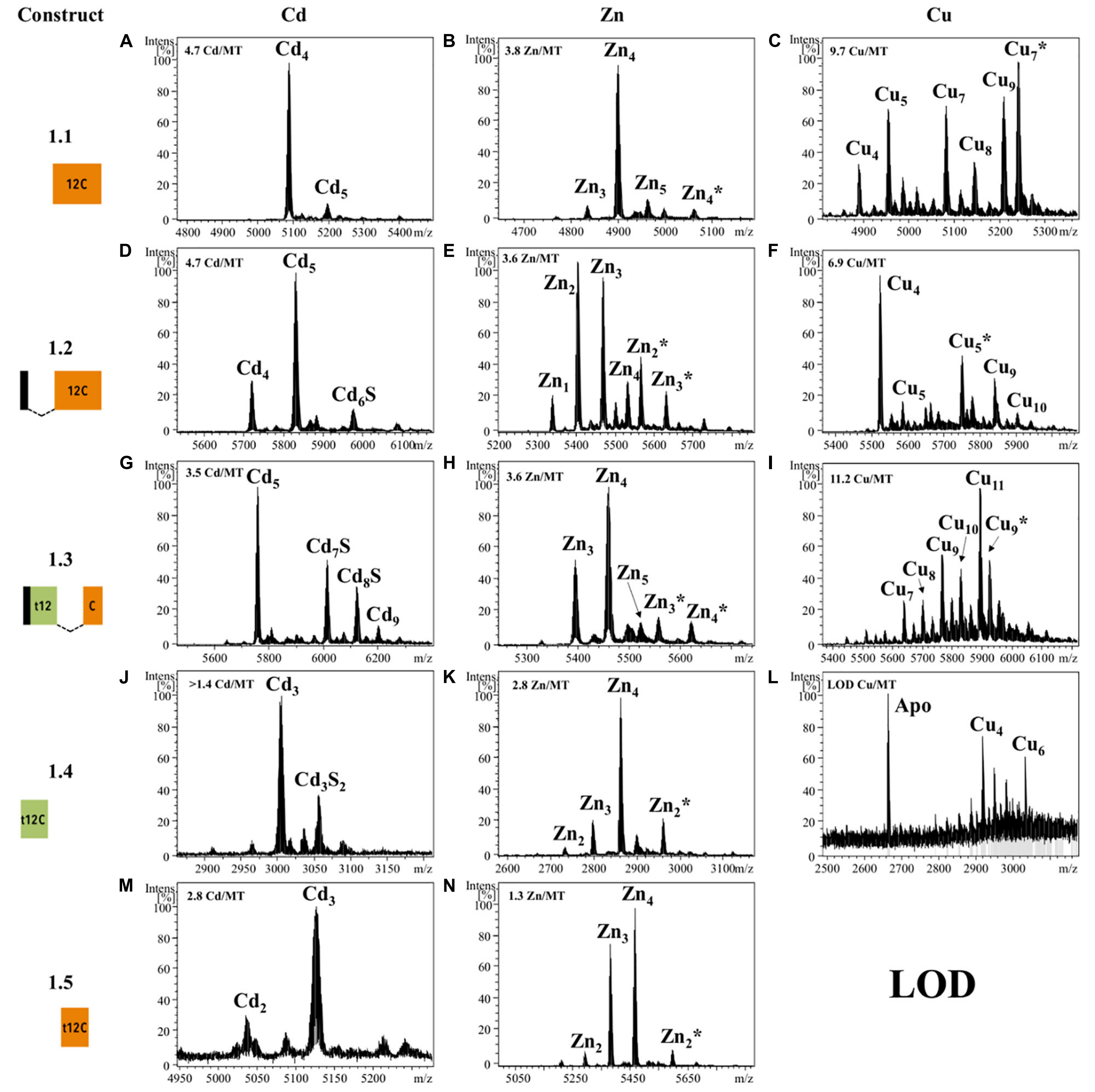


FIGURE 2 | Deconvoluted ESI-MS spectra recorded at pH 7.0 of *OdiMT1* constructs 1.1 **(A–C)**, 1.2 **(D–F)**, 1.3 **(G–I)**, 1.4 **(J–L)**, and 1.5 **(M,N)** bioproduced with Cd **(A,D,G,J,M)**, Zn **(B,E,H,K,N)**, or Cu **(C,F,I,L)**. *Y*-axis represents the relative intensities of the peaks while the *X*-axis represents the m/z ratio equivalent to the MW (Da). The metal-to-protein ratio obtained by ICP-AES is included on the top left of each panel. It is indicated those cases in which the sample concentration was lower than the limit of detection (LOD). Glycosylated species are marked with an asterisk (*).

physiological capabilities of the organisms to adapt to diverse conditions of metal bioavailability and other environmental stress situations.

Genetic Variation in *O. dioica* MTs

The peculiar structural and evolutionary characteristics of OdiMTs together with the high evolutionary rate of *O. dioica* (Denoeud et al., 2010; Berna et al., 2012; Berna and Alvarez-Valin, 2014), and significant level of sequence variation detected between *O. dioica* populations (Wang et al., 2015, 2020; Bliznina et al., 2021), prompted us to investigate the OdiMT sequences in several *O. dioica* populations worldwide distributed. We analyzed the MT sequences from animals from Norway (NOR), Oregon (ORE), Japan (Japanese specimens were from two different localizations, Osaka (OSA) and Okinawa (OKI), and we have analyzed

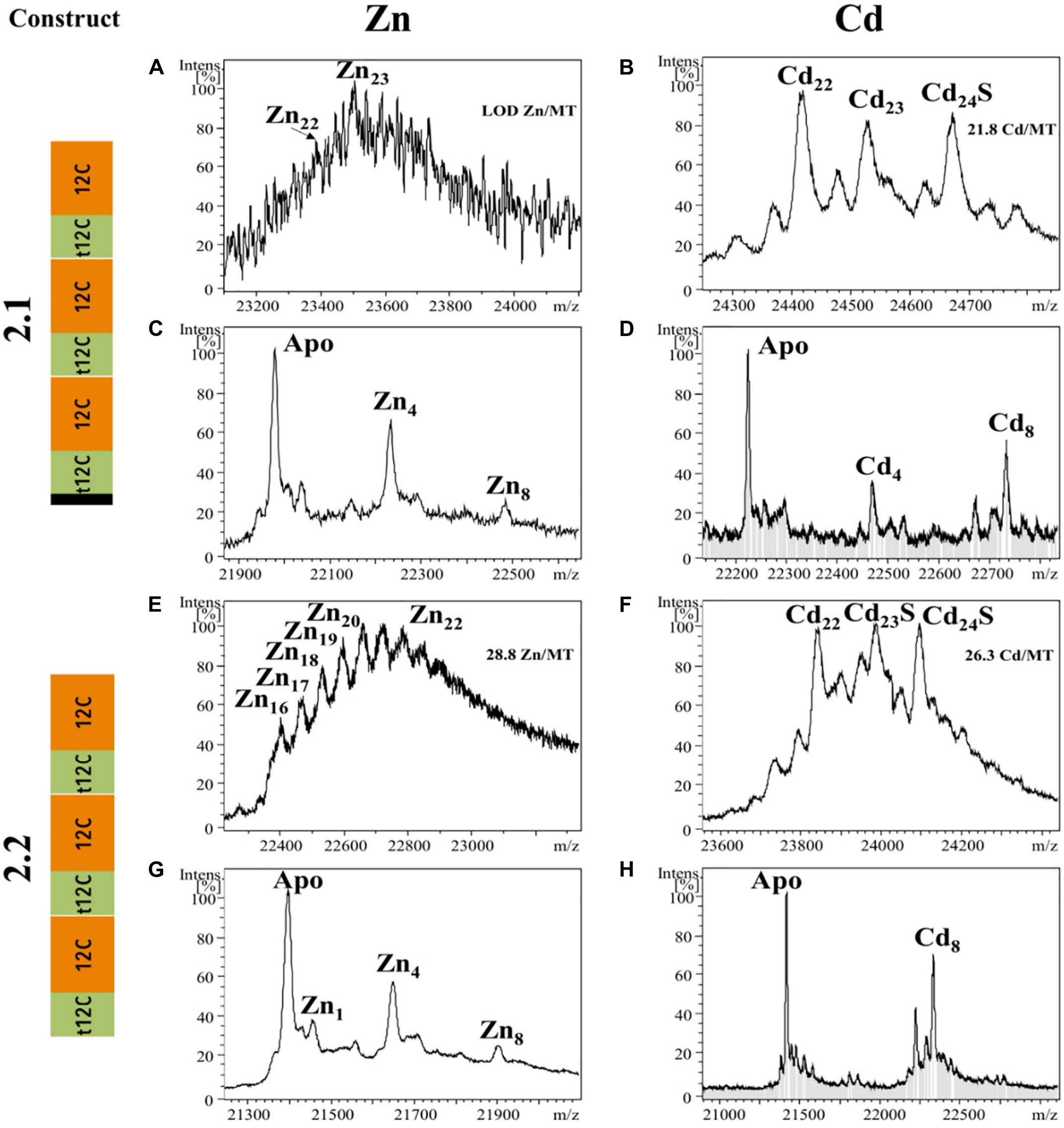


FIGURE 3 | Deconvoluted ESI-MS spectra recorded at pH 7.0 **(A,B,E,F)** and pH 2.4 **(C,D,G,H)** of *OdiMT2* constructs 2.1 **(A–D)** and 2.2 **(E–H)** bioproduced with Zn **(A,C,E,G)** or Cd **(B,D,F,H)**. *Y*-axis represents the relative intensities of the peaks while the *X*-axis represents the m/z ratio equivalent to the MW (Da). The metal-to-protein ratio obtained by ICP-AES is included on the top right of panels **(A,B,E,F)**. It is indicated those cases in which the sample concentration was lower than the limit of detection (LOD).

them separately), and Catalonia (CAT), representing four geographically distant *O. dioica* populations: north Atlantic, eastern Pacific, western Pacific and Mediterranean populations, respectively. We used as guiding reference the Norwegian sequences (*OdiMT*$_{\mathrm{NOR}}$) (**Figure 4** and **Supplementary Figure 1**) because they were the first sequences identified (Calatayud et al., 2018).

Both *OdiMT1* and *OdiMT2* genes were present in all analyzed populations (**Figure 3**), but comparison among populations revealed important differences affecting three aspect: (i) sequence variability, (ii) presence of non-functional allelic variants, and (iii) differences in the number of RU in the case of OdiMT2. Regarding sequence variability, protein sequence comparisons reveal amino-acid identities ranging from 94.4% between Norway and Barcelona, up to 63.9% when compared with the sequence of Okinawa, which appeared as the more distant population to any other one (**Table 2**). In contrast to the overall sequence variation – 82.6 and 87.7% of average amino acid identity (excluding the divergent Okinawa sequences) for OdiMT1 and OdiMT2, respectively, Cys residues were nearly invariants, with 98.3% preservation. This high Cys conservation contrasted with the fact that Cys are the less conserved amino acids in *O. dioica* proteins (Berna et al., 2012; Berná and Alvarez-Valin, 2015), suggesting that Cys substitutions are negatively selected due to the functional restrictions imposed by metal coordination.

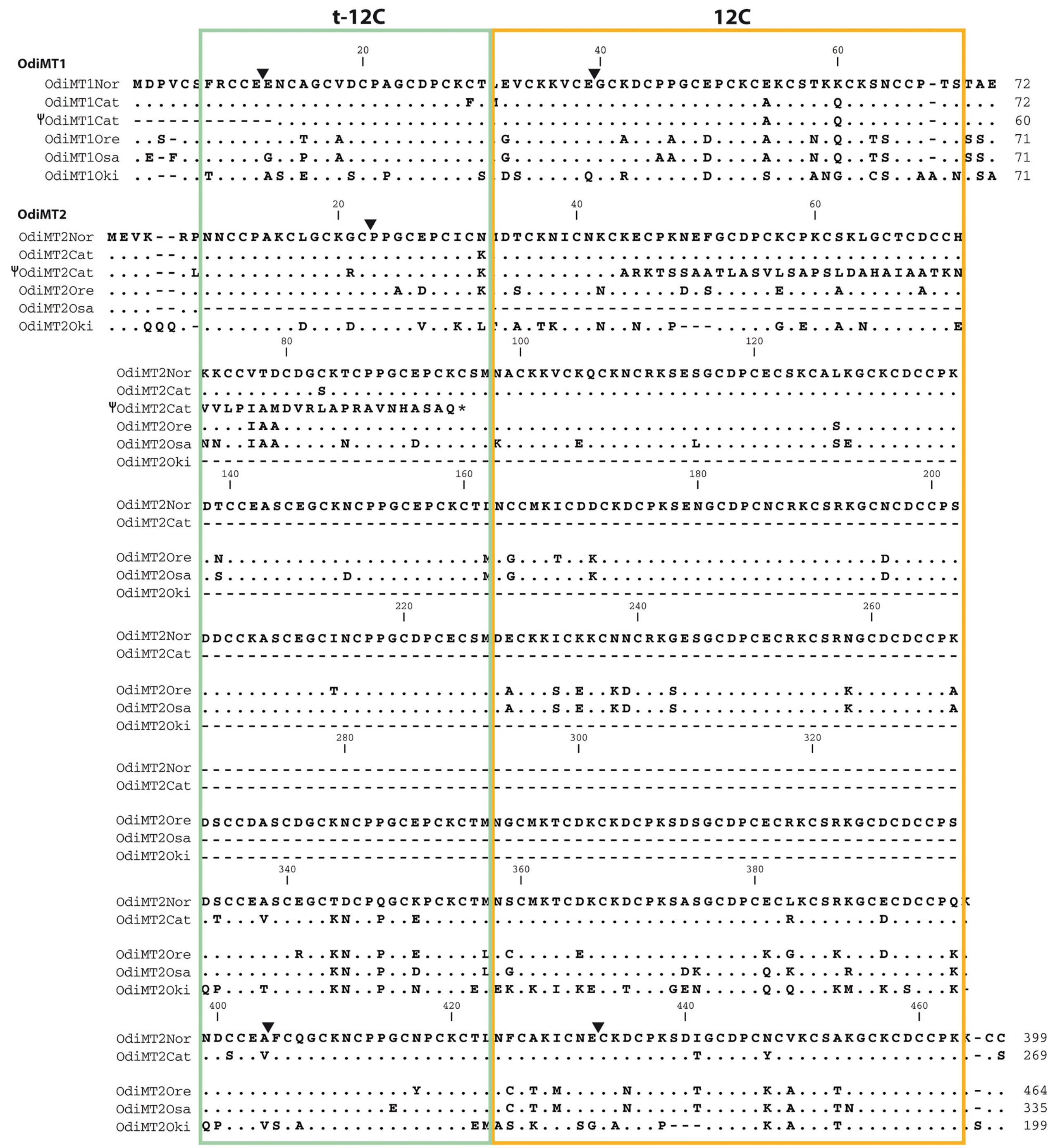

t-12C
12C
OdiMT1
OdiMT1Nor
OdiMT1Cat
ΨOdiMT1Cat
OdiMT1Ore
OdiMT1Osa
OdiMT1Oki
OdiMT2
OdiMT2Nor
OdiMT2Cat
ΨOdiMT2Cat
OdiMT2Ore
OdiMT2Osa
OdiMT2Oki

FIGURE 4 | Amino acid alignments of OdiMT1 and OdiMT2 variants form Norwegian (NOR), Catalonian (CAT), Oregonian (ORE), Osaka (OSA), and Okinawa (OKI) *O. dioica* populations. Norwegian sequences are used as reference; dots and dashes indicate amino acid identity and gaps, respectively. Black arrowheads indicate the intron positions relative to the amino acid sequences. The sequences of two hypothetical OdiMT pseudogenes –i.e., Ψ OdiMT1$_{CAT}$, lacking the first 12 residues, and Ψ OdiMT2$_{CAT}$, with a premature stop codon (*)– are also included. Trimmed 12-Cys domains (green box) and full-length 12-Cys domains (orange box) at the N- and C-terminal regions of the RU, respectively, are depicted.

Moreover, the fact that amino acid identities persistently were slightly lower than nucleotide identities (**Table 2**) indicated that nucleotide substitutions are significantly affecting non-synonymous positions, which can be considered an indication that MT variability among populations might probably be under positive selection.

Noteworthy, during the identification of the Catalonian MTs, we detected some non-functional allelic variants of *OdiMT1*

TABLE 2 | Percentage identity[a] of nucleotide (above diagonal) and amino acid sequences (below diagonal) of the OdiMT coding regions from five different *O. dioica* populations, Norway (NOR), Catalonia (CAT), Osaka (OSA), Okinawa (OKI), and Oregon (ORE).

	OdiMT1$_{\mathrm{NOR}}$	**OdiMT1$_{\mathrm{CAT}}$**	**OdiMT1$_{\mathrm{OSA}}$**	**OdiMT1$_{\mathrm{OKI}}$**	**OdiMT1$_{\mathrm{ORE}}$**
OdiMT1$_{\mathrm{NOR}}$	–	97.26	84.02	73.87	83.56
OdiMT1$_{\mathrm{CAT}}$	94.44	–	84.47	72.97	83.56
OdiMT1$_{\mathrm{OSA}}$	76.38	76.38	–	70.58	93.52
OdiMT1$_{\mathrm{OKI}}$	67.12	64.32	63.88	–	69.09
OdiMT1$_{\mathrm{ORE}}$	79.16	79.16	90.27	66.66	–
	OdiMT2$_{\mathrm{NOR}}$	**OdiMT2$_{\mathrm{CAT}}$**	**OdiMT2$_{\mathrm{OSA}}$**	**OdiMT2$_{\mathrm{OKI}}$**	**OdiMT2$_{\mathrm{ORE}}$**
OdiMT2$_{\mathrm{NOR}}$	–	96.42	86.62	70.40	86.68
OdiMT2$_{\mathrm{CAT}}$	94.07	–	85.81	70.72	87.42
OdiMT2$_{\mathrm{OSA}}$	85.42	81.16	–	73.95	92.77
OdiMT2$_{\mathrm{OKI}}$	67.79	68.75	67.59	–	73.76
OdiMT2$_{\mathrm{ORE}}$	87.00	86.67	91.69	69.23	–

[a] *For the calculation of sequence identities, a gap between two sequences has been considered as a single difference.*

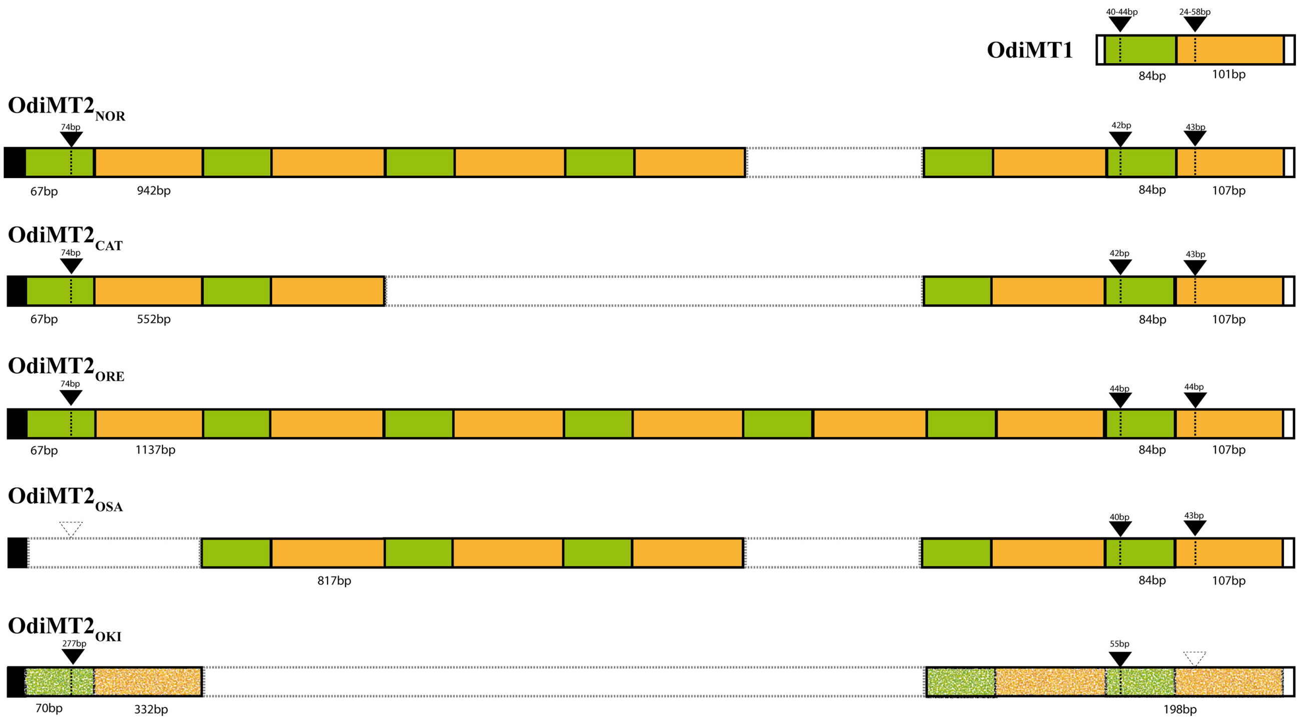


FIGURE 5 | Schematic representation of the organization of RU in OdiMTs. The t-12C/12C RU in OdiMT1s (top) and in OdiMT2 variants with six (OdiMT2$_{\mathrm{NOR}}$), four (OdiMT2$_{\mathrm{CAT}}$), seven (OdiMT2$_{\mathrm{ORE}}$), five (OdiMT2$_{\mathrm{OSA}}$), and three (OdiMT2$_{\mathrm{OKI}}$) RU are depicted following the color code of **Figure 1**. Repeats are positioned on the basis of intron positions (black arrowheads; dashed arrowheads denote absent introns) and phylogenetic analysis (**Supplementary Figure 2**). Exon and intron sizes are indicated. The high divergence of the RU of OdiMT2$_{\mathrm{OKI}}$ is highlighted by the dotted pattern.

and *OdiMT2* genes. One variant displayed a 35 nucleotide (nt) deletion at the 5'-end of the *OdiMT1* genomic region that eliminated the first exon, and therefore likely resulted in a pseudogene, Ψ-*OdiMT1*$_{CAT}$ (**Figure 4** and **Supplementary Figure 1**). In another variant, we found a 7 nt deletion at the beginning of the exon 2 of the *OdiMT2* that caused a frame-shift mutation that resulted in a change of the amino acid sequence and a premature stop codon (**Figure 4** and **Supplementary Figure 1**). The functionality of the truncated protein codified by Ψ-*OdiMT2*$_{CAT}$ was unlikely because only 44% (41 out of 92) of the amino acids were similar to the OdiMT2 sequence (**Figure 4**). The presence of these non-functional alleles for both MTs, Ψ-*OdiMT1*$_{CAT}$, and Ψ-*OdiMT2*$_{CAT}$, in the Mediterranean population opened the possibility that some *O. dioica* specimens might lack functional MTs. Further analyses of the presence and frequencies of non-functional variants in different populations could reveal different selective pressures related to variations in heavy-metal amounts of different environments.

The most conspicuous difference of *OdiMT* genes between *O. dioica* populations was the identification of *OdiMT2* encoding proteins with variable number of RUs (i.e., t-12C/12 pairs) (**Figures 4**, **5**). OdiMT2_{ORE} with seven repeats (RU_1–RU_7) was the longest one, followed by OdiMT2_{NOR}, with six repeats, OdiMT2_{OSA}, with five repeats, OdiMT2_{CAT}, with four repeats, and OdiMT2_{OKI}, with three repeats. The conservation in the *OdiMT2* alleles of the introns both at the 5′- and 3′-ends of the gene (**Figure 5**), together with phylogenetic reconstructions based on the sequence alignments of the repeats (**Supplementary Figure 2**), suggested that the increase/decrease in the number of repeats was the result of internal tandem duplications/losses due to recombination events. Phylogenetic analysis suggested that some expansions of the RU preceded the divergence of populations, while some independent gains or losses also occurred (**Supplementary Figure 2**). In OdiMT2_{OKI}, for instance, duplications of RU seemed to have taken place independently form other OdiMT2s, suggesting again that this population is the most divergent among all populations. Finally, the fact that in all variants the repeat units comprised complete t-12C/12C pairs supported a modular, step-wise evolution for OdiMTs, and agreed with the functional analysis of the domains that demonstrated that despite the autonomy of the 12C domain, the t-12C/12C pair appeared as an improved functional unit for divalent metal binding (see section "Functional Analysis of the Modularity in OdiMTs").

Proteins with domain repeats have been observed to be particularly common in multicellular species (Apic et al., 2001), especially in vertebrates (Björklund et al., 2006). The exact mechanism for repeat expansion remains to be discovered, but evidence of the expansion of repeats come from the fact that orthologous proteins might have different numbers of domain repeats in different species (Björklund et al., 2006). The number of domain repeats is, however, well conserved within a species, and it does not present intraspecific variability in terms of repeat unit gains/losses (Schaper et al., 2014). The population variability of OdiMT2s here exposed is, therefore, surprising, and *O. dioica* challenges again the standard patterns of gene and protein evolution, opening a new perspective in the functional and structural evolution of domain repeat proteins.

In summary, our results suggested that the modular organization provides MTs with a high structural and functional plasticity and dynamism, as it demonstrates the detection of variants with variable number of t-12C/12C repeats. These features seem to have facilitated the creation of large multi-modular MTs with high cysteine content and a high capacity of metal binding. Large multi-modular MTs have been described in other organisms (Palacios et al., 2014; Iturbe-Espinoza et al., 2016; Dallinger et al., 2020; Calatayud et al., 2021a,b), some of them associated with biological adaptations (Palacios et al., 2014; Iturbe-Espinoza et al., 2016; Jenny et al., 2016; Baumann et al., 2017). These large MTs did not show, however, the complexity level reached by *O. dioica* proteins, nor the inter-population variability described here. It seems, therefore, that *O. dioica* would be exploring the limits of chordate MT evolvability since, although we still do not know if adaptive selection to different environmental conditions would be driving the changes in the number of the t-12C/12C repeats, the more genetic variation there is, the greater the capacity for adaptive evolution of a biological system.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RA was responsible for the project coordination. RA, CC, ÒP, and MC conceived and designed the experiments. SC, CC, and RA collected the MT sequences from databases and elaborated the evolutionary inferences. SC performed the cloning and recombinant synthesis of the analyzed proteins. MG-R performed the ICP-AES, CD, UV-vis, and ESI-MS experiments. MG-R, ÒP, and MC analyzed the metal-binding data. SC and MG-R were responsible of the figures and tables. All authors discussed the experimental results, were responsible for writing the manuscript, commented on the manuscript, and agreed to its final version.

FUNDING

RA was supported by BIO2015-67358-C2-1-P. CC was supported by BFU2016-80601-P and PID2019-110562GB-I00. MC and ÒP were supported by BIO2015-67358-C2-2-P from the Spanish Ministerio de Ciencia e Innovación. MC and ÒP were members of the "Grup de Recerca de la Generalitat de Catalunya," ref. 2017SGR-864, and RA and CC of ref. 2017SGR-1665. MG-R acknowledges to the UAB the PIF grant.

ACKNOWLEDGMENTS

We acknowledge the Centres Científics i Tecnològics (CCiT) de la Universitat de Barcelona (DNA sequencing) and the Servei d'Anàlisi Química (SAQ) de la Universitat Autònoma de Barcelona (ICP-AES, ESI-MS) for allocating instrument time. We also thank to Sebastian Artime for experimental support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.702688/full#supplementary-material

REFERENCES

Adamo, G. M., Lotti, M., Tamas, M. J., and Brocca, S. (2012). Amplification of the CUP1 gene is associated with evolution of copper tolerance in Saccharomyces cerevisiae. *Microbiology* 158, 2325–2335. doi: 10.1099/mic.0.058024-0

Apic, G., Gough, J., and Teichmann, S. A. (2001). Domain combinations in archaeal, eubacterial and eukaryotic proteomes11Edited by G. von Heijne. *J. Mole. Biol.* 310, 311–325. doi: 10.1006/jmbi.2001.4776

Artells, E., Palacios, O., Capdevila, M., and Atrian, S. (2013). Mammalian MT1 and MT2 metallothioneins differ in their metal binding abilities. *Metallomics* 5, 1397–1410. doi: 10.1039/c3mt00123g

Baumann, C., Beil, A., Jurt, S., Niederwanger, M., Palacios, O., Capdevila, M., et al. (2017). Structural Adaptation of a Protein to Increased Metal Stress: NMR Structure of a Marine Snail Metallothionein with an Additional Domain. *Angew. Chem. Int. Ed. Engl.* 56, 4617–4622. doi: 10.1002/anie.201611873

Beil, A., Jurt, S., Walser, R., Schonhut, T., Guntert, P., Palacios, O., et al. (2019). The Solution Structure and Dynamics of Cd-Metallothionein from Helix pomatia Reveal Optimization for Binding Cd over Zn. *Biochemistry* 58, 4570–4581. doi: 10.1021/acs.biochem.9b00830

Berna, L., and Alvarez-Valin, F. (2014). Evolutionary genomics of fast evolving tunicates. *Genome Biol. Evol.* 6, 1724–1738. doi: 10.1093/gbe/evu122

Berná, L., and Alvarez-Valin, F. (2015). Evolutionary volatile Cysteines and protein disorder in the fast evolving tunicate *Oikopleura dioica*. *Mar. Genomics* 24, 47–54. doi: 10.1016/j.margen.2015.07.007

Berna, L., D'Onofrio, G., and Alvarez-Valin, F. (2012). Peculiar patterns of amino acid substitution and conservation in the fast evolving tunicate Oikopleura dioica. *Mol. Phylogenet. Evol.* 62, 708–717. doi: 10.1016/j.ympev.2011.11.013

Björklund, A. K., Ekman, D., and Elofsson, A. (2006). Expansion of protein domain repeats. *PLoS Comput. Biol.* 2:e114–e114. doi: 10.1371/journal.pcbi.0020114

Björklund, Åk, Ekman, D., Light, S., Frey-Skött, J., and Elofsson, A. (2005). Domain Rearrangements in Protein Evolution. *J. Mole. Biol.* 353, 911–923. doi: 10.1016/j.jmb.2005.08.067

Blindauer, C. A. (2014). "Metallothioneins," in *Binding, Transport and Storage of Metal Ions in Biological Cells*, eds W. Maret and A. Wedd (Cambridge, UK: The Royal Society of Chemistry), 594–653.

Bliznina, A., Masunaga, A., Mansfield, M. J., Tan, Y., Liu, A. W., West, C., et al. (2021). Telomere-to-telomere assembly of the genome of an individual Oikopleura dioica from Okinawa using Nanopore-based sequencing. *BMC Genomics* 22:222. doi: 10.1186/s12864-021-07512-6

Bongers, J., Walton, C. D., Richardson, D. E., and Bell, J. U. (1988). Micromolar protein concentrations and metalloprotein stoichiometries obtained by inductively coupled plasma atomic emission spectrometric determination of sulfur. *Anal. Chem.* 60, 2683–2686.

Braun, W., Wagner, G., Worgotter, E., Vasak, M., Kagi, J. H., and Wuthrich, K. (1986). Polypeptide fold in the two metal clusters of metallothionein-2 by nuclear magnetic resonance in solution. *J. Mol. Biol.* 187, 125–129. doi: 10.1016/0022-2836(86)90412-2

Calatayud, S., Garcia-Risco, M., Palacios, O., Capdevila, M., Cañestro, C., and Albalat, R. (2021a). Tunicates illuminate the enigmatic evolution of chordate metallothioneins by gene gains and losses, independent modular expansions and functional convergences. *Mole. Biol. Evol.* doi: 10.1093/molbev/msab184

Calatayud, S., Garcia-Risco, M., Pedrini-Martha, V., Eernisse, D. J., Dallinger, R., Capdevila, M., et al. (2021b). Modularity in Protein Evolution: Modular Organization and De Novo Domain Evolution in Mollusk Metallothioneins. *Mole. Biol. Evol.* 38, 424–436. doi: 10.1093/molbev/msaa230

Calatayud, S., Garcia-Risco, M., Rojas, N. S., Espinosa-Sanchez, L., Artime, S., Palacios, O., et al. (2018). Metallothioneins of the urochordate Oikopleura dioica have Cys-rich tandem repeats, large size and cadmium-binding preference. *Metallomics* 10, 1585–1594. doi: 10.1039/c8mt00177d

Capdevila, M., and Atrian, S. (2011). Metallothionein protein evolution: a miniassay. *J. Biol. Inorg. Chem.* 16, 977–989. doi: 10.1007/s00775-011-0798-3

Capdevila, M., Bofill, R., Palacios, Ò, and Atrian, S. (2012). State-of-the-art of metallothioneins at the beginning of the 21st century. *Coordinat. Chem. Rev.* 256, 46–62. doi: 10.1016/j.ccr.2011.07.006

Capdevila, M., Domènech, J., Pagani, A., Tío, L., Villarreal, L., and Atrian, S. (2005). Zn– and Cd–Metallothionein Recombinant Species from the Most Diverse Phyla May Contain Sulfide (S2−) Ligands. *Angew. Chem. Internat. Edn.* 44, 4618–4622. doi: 10.1002/anie.200501183

Catalan, A., Glaser-Schmitt, A., Argyridou, E., Duchen, P., and Parsch, J. (2016). An Indel Polymorphism in the MtnA 3′ Untranslated Region Is Associated with Gene Expression Variation and Local Adaptation in *Drosophila melanogaster*. *PLoS Genet.* 12:e1005987. doi: 10.1371/journal.pgen.1005987

Cols, N., Romero-Isart, N., Capdevila, M., Oliva, B., Gonzalez-Duarte, P., Gonzalez-Duarte, R., et al. (1997). Binding of excess cadmium(II) to Cd7-metallothionein from recombinant mouse Zn7-metallothionein 1. UV-VIS absorption and circular dichroism studies and theoretical location approach by surface accessibility analysis. *J. Inorg. Biochem.* 68, 157–166. doi: 10.1016/s0162-0134(97)00085-8

Costa, D., Marien, J., Janssens, T. K., van Gestel, C. A., Driessen, G., Sousa, J. P., et al. (2012). Influence of adaptive evolution of cadmium tolerance on neutral and functional genetic variation in Orchesella cincta. *Ecotoxicology* 21, 2078–2087. doi: 10.1007/s10646-012-0961-9

Dallinger, R., Zerbe, O., Baumann, C., Egger, B., Capdevila, M., Palacios, O., et al. (2020). Metallomics reveal a persisting impact of cadmium on the evolution of metal-selective snail metallothioneins. *Metallomics* 12, 702–720. doi: 10.1039/c9mt00259f

Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J. M., Da Silva, C., Brinkmann, H., et al. (2010). Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science* 330, 1381–1385. doi: 10.1126/science.1194167

Digilio, G., Bracco, C., Vergani, L., Botta, M., Osella, D., and Viarengo, A. (2009). The cadmium binding domains in the metallothionein isoform Cd7-MT10 from Mytilus galloprovincialis revealed by NMR spectroscopy. *JBIC J. Biol. Inorgan. Chem.* 14, 167–178. doi: 10.1007/s00775-008-0435-y

Dohmen, E., Klasberg, S., Bornberg-Bauer, E., Perrey, S., and Kemena, C. (2020). The modular nature of protein evolution: domain rearrangement rates across eukaryotic life. *BMC Evol. Biol.* 20:30. doi: 10.1186/s12862-020-1591-0

Edvardsen, R. B., Seo, H. C., Jensen, M. F., Mialon, A., Mikhaleva, J., Bjordal, M., et al. (2005). Remodelling of the homeobox gene complement in the tunicate Oikopleura dioica. *Curr. Biol.* 15, R12–R13.

Egli, D., Domènech, J., Selvaraj, A., Balamurugan, K., Hua, H., Capdevila, M., et al. (2006). The four members of the Drosophila metallothionein family exhibit distinct yet overlapping roles in heavy metal homeostasis and detoxification. *Genes Cells* 11, 647–658. doi: 10.1111/j.1365-2443.2006.00971.x

Ferrández-Roldán, A., Martí-Solans, J., Cañestro, C., and Albalat, R. (2019). "Oikopleura dioica: An Emergent Chordate Model to Study the Impact of Gene Loss on the Evolution of the Mechanisms of Development," in *Evo-Devo: Non-model Species in Cell and Developmental Biology*, eds W. Tworzydlo and S. M. Bilinski (Berlin: Springer International Publishing), 63–105.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Guirola, M., Perez-Rafael, S., Capdevila, M., Palacios, O., and Atrian, S. (2012). Metal dealing at the origin of the Chordata phylum: the metallothionein system and metal overload response in amphioxus. *PLoS One* 7:e43299. doi: 10.1371/journal.pone.0043299

Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci.* 89:10915. doi: 10.1073/pnas.89.22.10915

Iturbe-Espinoza, P., Gil-Moreno, S., Lin, W., Calatayud, S., Palacios, O., Capdevila, M., et al. (2016). The Fungus Tremella mesenterica Encodes the Longest Metallothionein Currently Known: Gene, Protein and Metal Binding Characterization. *PLoS One* 11:e0148651. doi: 10.1371/journal.pone.0148651

Jenny, M. J., Payton, S. L., Baltzegar, D. A., and Lozier, J. D. (2016). Phylogenetic Analysis of Molluscan Metallothioneins: Evolutionary Insight from *Crassostrea virginica*. *J. Mol. Evol.* 83, 110–125. doi: 10.1007/s00239-016-9758-4

Kojima, Y., Binz, P.-A., and Kägi, J. H. R. (1999). "Nomenclature of metallothionein: Proposal for a revision," in *Metallothionein IV. Basel: Birkhäuser Basel*, ed. C. D. Klaassen (Basel: Birkhäuser Basel), 3–6.

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531

Marino, S. M., and Gladyshev, V. N. (2010). Cysteine Function Governs Its Conservation and Degeneration and Restricts Its Utilization on Protein Surfaces. *J. Mole. Biol.* 404, 902–916. doi: 10.1016/j.jmb.2010.09.027

Martí-Solans, J., Ferrández-Roldán, A., Godoy-Marín, H., Badia-Ramentol, J., Torres-Aguila, N. P., Rodríguez-Marí, A., et al. (2015). Oikopleura dioica culturing made easy: a low-cost facility for an emerging animal model in EvoDevo. *Genesis* 53, 183–193. doi: 10.1002/dvg.22800

Munoz, A., Forsterling, F. H., Shaw, C. F. III, and Petering, D. H. (2002). Structure of the (113)Cd(3)beta domains from Homarus americanus metallothionein-1: hydrogen bonding and solvent accessibility of sulfur atoms. *J. Biol. Inorg. Chem.* 7, 713–724. doi: 10.1007/s00775-002-0345-3

Nam, Y. K., and Kim, E. J. (2017). Diversification and domain evolution of molluskan metallothioneins: a mini review. *Fish. Aquatic Sci.* 20:8. doi: 10.1186/s41240-017-0054-z

Otvos, J. D., and Armitage, I. M. (1980). Structure of the metal clusters in rabbit liver metallothionein. *Proc. Natl. Acad. Sci.* 77:7094. doi: 10.1073/pnas.77.12.7094

Palacios, O., Atrian, S., and Capdevila, M. (2011). Zn- and Cu-thioneins: a functional classification for metallothioneins? *J. Biol. Inorg. Chem.* 16, 991–1009. doi: 10.1007/s00775-011-0827-2

Palacios, O., Espart, A., Espin, J., Ding, C., Thiele, D. J., Atrian, S., et al. (2014). Full characterization of the Cu-, Zn-, and Cd-binding properties of CnMT1 and CnMT2, two metallothioneins of the pathogenic fungus Cryptococcus neoformans acting as virulence factors. *Metallomics* 6, 279–291. doi: 10.1039/c3mt00266g

Riek, R., Precheur, B., Wang, Y., Mackay, E. A., Wider, G., Guntert, P., et al. (1999). NMR structure of the sea urchin (Strongylocentrotus purpuratus) metallothionein MTA. *J. Mol. Biol.* 291, 417–428. doi: 10.1006/jmbi.1999.2967

Schaper, E., Gascuel, O., and Anisimova, M. (2014). Deep Conservation of Human Protein Tandem Repeats within the Eukaryotes. *Mole. Biol. Evol.* 31, 1132–1148. doi: 10.1093/molbev/msu062

Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vašák, M., Kägi, J. H. R., et al. (1988). Conformation of [Cd7]-metallothionein-2 from rat liver in aqueous solution determined by nuclear magnetic resonance spectroscopy. *J. Mole. Biol.* 203, 251–268. doi: 10.1016/0022-2836(88)90106-4

Stillman, M. J., Cai, W., and Zelazowski, A. J. (1987). Cadmium binding to metallothioneins. Domain specificity in reactions of alpha and beta fragments, apometallothionein, and zinc metallothionein with Cd2+. *J. Biol. Chem.* 262, 4538–4548. doi: 10.1016/S0021-9258(18)61226-8

Tomas, M., Domènech, J., Capdevila, M., Bofill, R., and Atrian, S. (2013). The sea urchin metallothionein system: Comparative evaluation of the SpMTA and SpMTB metal-binding preferences. *FEBS Open Bio* 3, 89–100. doi: 10.1016/j.fob.2013.01.005

Vasak, M., and Meloni, G. (2011). Chemistry and biology of mammalian metallothioneins. *J. Biol. Inorg. Chem.* 16, 1067–1078. doi: 10.1007/s00775-011-0799-2

Wang, K., Omotezako, T., Kishi, K., Nishida, H., and Onuma, T. A. (2015). Maternal and zygotic transcriptomes in the appendicularian, Oikopleura dioica: novel protein-encoding genes, intra-species sequence variations, and trans-spliced RNA leader. *Dev. Genes Evol.* 225, 149–159. doi: 10.1007/s00427-015-0502-7

Wang, K., Tomura, R., Chen, W., Kiyooka, M., Ishizaki, H., Aizu, T., et al. (2020). A genome database for a Japanese population of the larvacean Oikopleura dioica. *Dev. Growth Differ.* 62, 450–461. doi: 10.1111/dgd.12689

Conflict of Interest: 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 © 2021 Calatayud, Garcia-Risco, Capdevila, Cañestro, Palacios and Albalat. 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) and the copyright owner(s) 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 Cell and Developmental Biology

MINI REVIEW
published: 06 July 2021
doi: 10.3389/fcell.2021.702787

Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease

Juan J. Tena and José M. Santos-Pereira**

Centro Andaluz de Biología del Desarrollo, Consejo Superior de Investigaciones Científicas/Universidad Pablo de Olavide, Seville, Spain

Animal genomes are folded in topologically associating domains (TADs) that have been linked to the regulation of the genes they contain by constraining regulatory interactions between *cis*-regulatory elements and promoters. Therefore, TADs are proposed as structural scaffolds for the establishment of regulatory landscapes (RLs). In this review, we discuss recent advances in the connection between TADs and gene regulation, their relationship with gene RLs and their dynamics during development and differentiation. Moreover, we describe how restructuring TADs may lead to pathological conditions, which explains their high evolutionary conservation, but at the same time it provides a substrate for the emergence of evolutionary innovations that lay at the origin of vertebrates and other phylogenetic clades.

Keywords: topologically associating domains, evolution, chromatin structure, genetic diseases, evolutionary novelties, regulatory landscapes

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Jonas Ibn-Salem,
Translationale Onkologie an der Universitätsmedizin der Johannes Gutenberg-Universität Mainz, Germany
Ann Campbell Burke,
Wesleyan University, United States
Alvaro Rada-Iglesias,
Instituto de Biomedicina y Biotecnología de Cantabria, Consejo Superior de Investigaciones Científicas, Spain

****Correspondence:***
Juan J. Tena
jjtenagu@upo.es
José M. Santos-Pereira
jmsanper1@upo.es

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *29 April 2021*
Accepted: *17 June 2021*
Published: *06 July 2021*

Citation:
Tena JJ and Santos-Pereira JM (2021) Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease. Front. Cell Dev. Biol. 9:702787. doi: 10.3389/fcell.2021.702787

INTRODUCTION

In the last years, the development of chromosome conformation capture (3C) techniques, together with remarkable advances in live cell imaging, have expanded our knowledge about the structural organization of animal genomes (Lieberman-Aiden et al., 2009; McCord et al., 2020; Jerkovic' and Cavalli, 2021). 3C techniques consist of restriction enzyme digestion of crosslinked chromatin, followed by proximity ligation to generate chimeric molecules that are interpreted as interactions between two genomic regions (McCord et al., 2020). HiC is the genome-wide version of 3C techniques and its increasing resolution has allowed to discern different levels of 3D folding at different scales. At the megabases scale, gene-rich transcriptionally active regions tend to interact among them, while gene-poor heterochromatic regions also interact more frequently, leading to A and B compartments, respectively. At the sub-megabase scale, chromatin domains with high interaction frequency and relatively isolated from neighbor regions form topologically associating domains (TADs). Finally, below the scale of TADs, chromatin loops are formed by strong interactions between specific genomic regions, i.e., CCCTC-binding factor (CTCF) and enhancer-promoter loops (Bonev and Cavalli, 2016; Rada-Iglesias et al., 2018; Rowley and Corces, 2018).

Topologically associating domains are believed to facilitate interactions between *cis*-regulatory elements (CREs) and their target promoters, which otherwise would not occur at enough frequency to ensure a robust target gene expression (Galupa and Heard, 2017; Franke and Gomez-Skarmeta, 2018; Furlong and Levine, 2018). Therefore, TADs have been proposed as structural scaffolds for regulatory landscapes (RLs; Acemel et al., 2017), which are defined as "large genomic regions containing several long-range-acting regulatory sequences that control one or several target genes in a coordinated manner" (Spitz et al., 2003; Bolt and Duboule, 2020). However, whether TADs represent a privileged functional level in the chromosome folding hierarchy has been challenged by the increasing resolution of HiC assays that have uncovered nested structures at the subTAD level with relative insulation among them (Rao et al., 2014; Zhan et al., 2017; Hsieh et al., 2020; Krietenstein et al., 2020). In this review, we discuss the connections of the 3D genome with gene expression, the relationship between TADs and RLs, and their dynamics in the context of development, focusing on disease and evolution.

FORMATION OF TADs BY LOOP EXTRUSION

TAD boundaries are enriched for the binding of CTCF (Phillips and Corces, 2009; Dixon et al., 2012; Merkenschlager and Nora, 2016), an 11-zinc-finger DNA binding protein that was previously known by its role in transcriptional insulation (Filippova et al., 1996; Bell et al., 1999). CTCF is distributed throughout the genome not only at TAD boundaries, but also at many other sites. However, its binding at boundaries has preference for a specific orientation, with its DNA binding motifs often positioned in convergent orientation between the two boundaries that define a TAD (Rao et al., 2014). CTCF co-localizes and determines the presence at TAD boundaries of the cohesin complex, which is also involved in establishing chromosomal interactions (Parelho et al., 2008; Hadjur et al., 2009). These latter observations led to the proposal of the so called "loop extrusion" model for TAD formation involving cohesin and CTCF (de Wit et al., 2015; Sanborn et al., 2015; Fudenberg et al., 2016). According to this model, TADs would arise by chromatin extrusion mediated by cohesin and counteracted by CTCF-mediated insulation, thus explaining both the increased interaction frequency within TADs and their relative insulation from neighbor TADs. Indeed, a recent study has shown that the N-terminal domain of CTCF is essential for blocking cohesin translocation from the interior of TADs, providing a molecular basis for the requirement of a specific binding polarity of CTCF for chromosome folding (Nora et al., 2020). The predictions of this model of TAD formation have been corroborated by the acute depletion of cohesin and CTCF in mammalian cells, which shows loss of chromatin loops and TAD insulation (Nora et al., 2017; Rao et al., 2017), as well as the depletion of factors regulating cohesin loading on chromatin, which leads to differences in the length of loops formed (Busslinger et al., 2017; Haarhuis et al., 2017; Schwarzer et al., 2017; Wutz et al., 2017).

REGULATORY LANDSCAPES AND CHROMATIN STRUCTURE

Regulatory landscapes contain CREs that control the expression of their target genes (Bolt and Duboule, 2020). Different mechanisms have been proposed to explain the transcriptional control mediated by CREs: *tracking*, which implies that the RNA polymerase II would bind to enhancers and track along chromatin synthesizing RNA until it reaches the promoter; *linking*, where transcription factors (TFs) would oligomerize from the enhancer to the promoter; or *looping*, in which factors bound to both sides of the loop (enhancer and promoter) would interact with each other (Furlong and Levine, 2018). The latter mechanism is widely accepted and would be favored by CTCF- and cohesin-dependent chromatin folding and promoted by mediator and TFs as Yin Yang 1, a zinc-finger DNA binding protein that form dimers similarly to CTCF and anchors enhancer-promoter interactions (Kagey et al., 2010; Dowen et al., 2014; Weintraub et al., 2017). Direct evidence of the functionality of enhancer-promoter looping comes from studies in the β-globin locus, where forcing these interactions by artificial zinc fingers leads to gene activation (Deng et al., 2014).

An interesting feature of vertebrate RLs is that enhancers are broadly distributed throughout them and not in a gene-centric manner (Symmons et al., 2014). Indeed, the RLs of developmental genes, such as those encoding for lineage-specific TFs or signaling molecules, are characterized by their large sizes and the abundance of enhancers that confer tissue-specific expression to their target genes, thus explaining their common pleiotropy (Bolt and Duboule, 2020). How do enhancers regulate only their target genes? This may be explained in part by the coincidence of RLs coordinates with those of TADs (Symmons et al., 2014). According to this observation, RLs would be confined within TADs and TAD boundaries would correspond to transitions of regulatory domains, therefore ensuring the contact of enhancers with the appropriate target genes and avoiding promiscuous interference with genes located in neighbor domains (see below). However, restriction of RLs within TADs is not enough for gene activation and other mechanisms may influence the outcome of enhancer-promoter contacts (Serebreni and Stark, 2021), including phase separation, which refers to local microenvironments resulting from weak multivalent interactions that concentrate some factors and exclude others (Banani et al., 2017), and enhancer-promoter compatibility (see below).

The RL of the *Shh* gene, encoding an important morphogen involved in the patterning of the developing neural tube and limb buds, is one of the best studied cases in vertebrates. This gene is located within an evolutionary conserved TAD of around 1 Mb in size that contains other non-related genes and multiple tissue-specific enhancers (Lettice et al., 2003; Sagai et al., 2005, 2009; Jeong et al., 2006). In particular, the enhancer known as ZRS [zone of polarizing activity (ZPA) regulatory sequence] is responsible for *Shh* expression in the limb bud and is located around 900-kb away, in the intron 5 of the *Lmbr1* gene (Lettice et al., 2003). Disruption of the ZRS sequence in mice causes loss of limbs, while in humans point mutations cause preaxial polydactyly (Sagai et al., 2005). Other paradigmatic examples of vertebrate RLs are the loci containing the *HoxD* and *HoxA* gene clusters, homeobox genes that are largely responsible for the patterning of several body structures including limbs. These clusters are located between two adjacent TADs that compartmentalize long-range regulatory interactions in two blocks at the spatial and temporal levels

(Lonfat and Duboule, 2015). Thus, enhancers in the 3′ TAD preferentially contact "anterior" *Hox* genes, while enhancers in the 5′ TAD mostly interact with "posterior" *Hox* genes (Lonfat et al., 2014). In addition, a switch in the interactions of central genes explains a sequential transition between two regulatory phases (Andrey et al., 2013). The regulatory activity of these enhancers is therefore combined to generate the collinearity of *Hox* genes. Apart from these well-known examples, other cases have been studied and recently reviewed (Bolt and Duboule, 2020).

THE LINK BETWEEN CHROMATIN STRUCTURE AND FUNCTION

The relationship between TADs and gene expression remains an issue of open debate due to apparent discrepancies between different approaches. Studies analyzing structural variations encompassing TAD boundaries have provided strong links between the chromatin architecture of particular loci and the expression of the nearby developmental genes, which become miss-expressed causing developmental abnormalities (Spielmann et al., 2018; Ibrahim and Mundlos, 2020). Although the adoption of enhancers by genes that were not previously in contact with them explains some of these phenotypes (Lupiáñez et al., 2015; Franke et al., 2016), this is not always the case (Laugsch et al., 2019) and the effects of structural variations on gene expression are context-specific (**Figure 1A**), which suggests that additional mechanisms may be involved. Indeed, fusion of neighbor TADs by boundary removal at the *Sox9-Kcnj2* locus does not result in major effects on gene expression (Despang et al., 2019). A study in *Drosophila* using highly rearranged balancer chromosomes concluded that TAD rearrangements did not result in altered expression of most genes (Ghavi-Helm et al., 2019). In agreement with this, the works describing alterations in chromatin structure and boundary deletions in the *Shh* locus show minor consequences on *Shh* expression not leading to developmental phenotypes (Paliou et al., 2019; Williamson et al., 2019).

The absence of transcriptional effects observed in some cases of TAD disruption may be explained by the lack of compatibility between enhancers and promoters. In this sense, it has been reported that transcriptional cofactors are able to activate only specific core promoters (Haberle et al., 2019), and that a particular group of highly conserved developmental enhancers, known as poised enhancers, contain CpG islands and are able to activate only developmental promoters also harboring CpG islands (Pachano et al., 2020). Moreover, it has been shown that enhancers within genomic regulatory blocks, which are regions containing a number of highly conserved non-coding elements, only activate particular target genes and not syntenic bystander genes (Akalin et al., 2009). While target genes commonly encode for developmental TFs and have promoters showing long CpG islands and multiple TF motifs, bystander genes often encode for proteins involved in unrelated functions, have different expression patterns and contain promoters with short CpG islands and few TF motifs. These observations indicate that enhancer-promoter compatibility mechanisms may determine the consequences of TAD restructuring at the transcriptional level.

On the other hand, several studies have attempted to remove architectural proteins and assess the effects of their loss genome wide. In this sense, the acute depletion of cohesin or CTCF in mammalian cells shows only a moderate effect in gene expression, affecting from some hundreds to few thousands of genes in different systems (Nora et al., 2017; Rao et al., 2017; Kubo et al., 2021). These studies have been limited *in vivo* due to the essential nature of cohesin and CTCF (Moore et al., 2012; Ju et al., 2013); however, this limitation was recently overcome by generating zebrafish *ctcf* knock-out embryos (Franke et al., 2020). In these embryos, a prolonged maternal contribution allows the survival of the mutant embryos until larval stages, when the absence of CTCF results in the miss-regulation of thousands of genes enriched in developmental functions (Franke et al., 2020). Although part of the effects seen in CTCF depletion or knock-out approaches may be indirect due to CTCF function as a TF, a subset of chromatin interactions involving lineage-specific genes change upon CTCF loss (Kubo et al., 2021). These data suggest that chromatin architecture might not be essential for the expression of most genes, but instead provide robustness to the expression of developmental genes that are frequently regulated by many CREs within complex RLs.

CHROMATIN STRUCTURE DYNAMICS DURING DEVELOPMENT AND DIFFERENTIATION

An important question in the field of chromatin structure is how variable are TADs during dynamic processes, including embryonic development or cell differentiation. Studies of chromatin structure during animal early development have revealed a phase in which TADs become undetectable followed by a progressive reestablishment of chromatin folding (van der Weide and de Wit, 2019). However, there are some differences between species regarding when this phase happens in relation to the zygotic genome activation (ZGA). While in mammals the period without detectable chromatin structure takes place before the ZGA, with structure being progressively reestablished from there (Du et al., 2017; Gassler et al., 2017; Chen et al., 2019), similar to *Drosophila* (Hug et al., 2017), in zebrafish and medaka embryos it occurs during the ZGA and persist after it, being reestablished mostly during gastrulation (Kaaij et al., 2018; Nakamura et al., 2018), but there are observed discrepancies regarding the existence of chromatin structure before the ZGA in zebrafish (Kaaij et al., 2018; Wike et al., 2021). Although once formed, TADs seem to be very stable at later developmental stages, little is known about the correlation of these structures with progressive changes in gene expression.

Several studies have reported that TADs are largely stable when compared between embryonic stem cells and differentiated cells (Dixon et al., 2012, 2015; Nora et al., 2012). Consistent with this observation, a recent study has shown that chromatin structure in *Drosophila* embryos is preserved in different

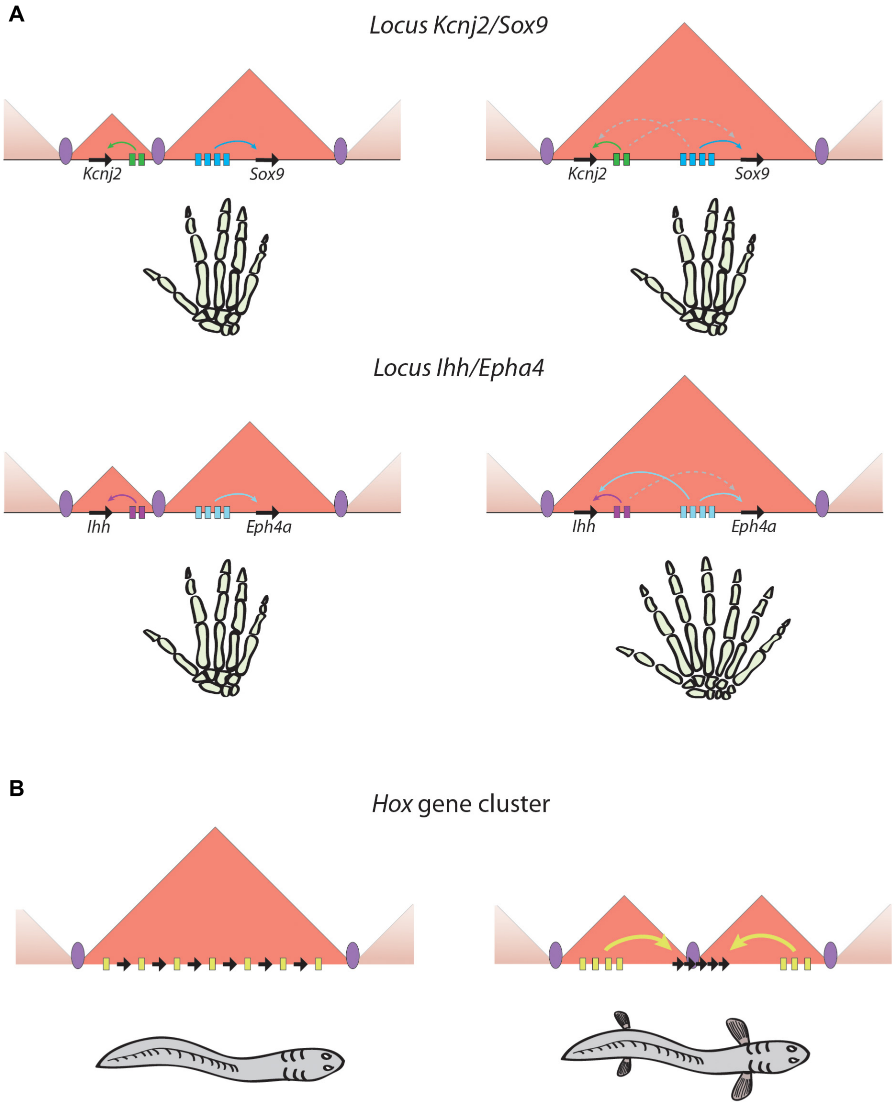


FIGURE 1 | Phenotypical outputs of regulatory landscapes rearrangements. **(A)** Similar rearrangements may lead to very different situations. In the two examples illustrated in this panel, the fusion of two adjacent TADs produces no detectable phenotype in the case of the *Kcnj2/Sox9 locus* (Despang et al., 2019), and a striking polydactylia phenotype in the case of the *Ihh/Eph4a locus* (Lupiáñez et al., 2015). **(B)** TAD restructuring can generate evolutionary innovations by gaining new regulatory inputs or by increasing regulatory complexity. An example is the splitting of the *HoxD* and *HoxA* ancestral TAD found in non-vertebrate chordates in two TADs that are vertebrate-specific and could favor the emergence of vertebrate appendages (Acemel et al., 2016). Black arrows, genes; colored rectangles, enhancers; colored arrows, productive enhancer-promoter interactions; dashed arrows, unproductive enhancer-promoter interactions; purple ellipses, TAD boundaries.

embryonic tissues despite lineage-specific gene expression (Ing-Simmons et al., 2021). However, another report showed a high variability of TAD boundaries across 37 human cell types (McArthur and Capra, 2021). It is worth noting that these apparent discrepancies might be explained either by the different biological systems used or by different resolutions in HiC experiments, as well as different computational methods to call TADs or TAD boundaries. At lower resolutions than TADs, it was shown that variations in intra-TAD contacts during mammalian differentiation corresponded with switches between active and inactive chromatin modifications and gene expression (Dixon et al., 2015). Indeed, enhancer-promoter interactions are highly dynamic and cell-type-specific during neural and erythroid differentiation, accompanying the activation of lineage-specific genes (Bonev et al., 2017; Oudelaar et al., 2020; Kubo et al., 2021). However, enhancer-promoter interactions are stably formed before gene activation in other contexts. Indeed, poised enhancers required for the activation of anterior neural genes are already engaged in contacts with their target genes in mESCs in a polycomb repressive complex 2-dependent manner (Cruz-Molina et al., 2017). However, enhancer-promoter loops in *Drosophila* precede target gene transcription and even TAD formation, being associated with paused RNA polymerase (Ghavi-Helm et al., 2014; Espinola et al., 2021). These observations suggest that the dynamics or stability of TADs and enhancer-promoter loops are highly context-specific, some of them being stable and others being dynamically regulated during development and differentiation.

ALTERATIONS OF THE 3D GENOME ASSOCIATED WITH DISEASE

Strong evidence for the importance of genome architecture for correct gene expression comes from studies showing that the disruption of 3D structure in particular loci, either by genomic rearrangements or alterations in TAD boundaries, lead to pathological situations, including developmental disorders or cancer (Ibrahim and Mundlos, 2020). A number of cases reported in the last years have provided a link between structural variations and disease. At the *Epha4* locus, deletions, duplications and inversions disrupting TAD structure cause several limb malformations, including brachydactyly, polydactyly and F-syndrome, due to *de novo* enhancer-promoter interactions that lead to gene miss-expression of the surrounding genes *Wnt6*, *Ihh* and *Pax3* (**Figure 1A**; Lupiáñez et al., 2015). At the *Sox9* locus, duplications encompassing the neighbor *Kcnj2* gene lead to the formation of a "neo-TAD" in which *Kcnj2* is miss-regulated by interactions with new enhancers, leading to a limb malformation phenotype known as Cooks syndrome (Franke et al., 2016). A similar situation has been reported in the locus of *GDPD1* gene in autosomal-dominant retinitis pigmentosa, caused by a rearrangement that place a TAD border and several retinal enhancers within this locus. This generates a "neo-TAD" and new contacts between the retinal enhancers and *GDPD1* gene, which is overexpressed and likely contributes to the disease (de Bruijn et al., 2020). Finally, inversions at the *TFAP2A* locus in branchiooculofacial syndrome patients have been shown to disconnect this gene from its neural crest-specific enhancers, leading to its reduced expression and explaining the patient's phenotype (Laugsch et al., 2019).

Genomic rearrangements can also lead to the fusion of genes that become overexpressed and function as oncogenes. This is the case of chromosomal translocations fusing the genes *PAX3* and *FOXO1* in alveolar rhabdomyosarcoma, which result in the fusion of both RLs and the activation of transcription from *PAX3* promoter by enhancers from the *FOXO1* RL (Vicente-Garcia et al., 2017). At larger scales, it has been found that higher order chromatin-folding structures can modulate interactions between different loci spanning several megabases in a highly aggressive type of squamous cell carcinoma (Alekseyenko et al., 2015; Rosencrance et al., 2020). The appearance of these interactions, called "megadomains," responds to the formation of large regions of hyperacetylated chromatin due to the BRD4-NUT fusion oncoprotein. These fusions are generated by genomic translocations, usually between genes like *BRD4* or *BRD3* and *NUT*, which recruits the histone acetyltransferase p300 leading to hyperacetylated regions of up to 2 Mb. Both intra- and interdomain interactions are up-regulated, as well as the expression of the oncogenes *SOX2*, *TP63* and *MYC*, which eventually contribute to tumorigenesis (Alekseyenko et al., 2015; Rosencrance et al., 2020; Eagen and French, 2021).

Apart from genomic rearrangements, alterations of TAD boundaries have also been linked to disease. The analyses of copy-number variants involving TAD boundary deletions revealed that a substantial proportion of cases could be explained by enhancer adoption (Ibn-Salem et al., 2014). In a recent work, CTCF binding sites surrounding the well-known *ZRS* enhancer of the *Shh* gene (see above) were removed by CRISPR-mediated genome editing in mice. Strikingly, the lack of these CTCF binding sites reduced the interaction between *ZRS* and *Shh* promoter, as well as the expression of the latter (Paliou et al., 2019). Nevertheless, this does not lead to a clear phenotype in these mice unless a hypomorphic allele of the ZRS is used, suggesting that chromatin structure provides robustness but does not determine enhancer-promoter communication. Indeed, a recent study from The Cancer Genome Atlas showed that only 14% of cancer-associated TAD boundary deletions resulted in significant changes in expression of the nearby genes (Akdemir et al., 2020). Moreover, mutations in CTCF binding sites have been frequently found in cancer (Katainen et al., 2015), which might lead to miss-expression of nearby genes causing tumorigenesis. Indeed, the disruption of boundaries demarcating insulated neighborhoods, which are chromatin domains smaller in size than TADs, leads to aberrant activation of proto-oncogenes, such as *TAL1* and *LMO2* associated with T-cell acute lymphoblastic leukemia (Hnisz et al., 2016). In the opposite situation, the overexpression of the Neurotensin gene *NTS*, a central nervous system neurotransmitter, has been recently related with melanomas due to a gained CTCF-mediated chromatin loop that establish contacts between the *NTS* promoter and a CRE in the intron of the *LRRIQ1* gene located 800 Kb away (Chai et al., 2021).

It is worth to remark that mutations in the coding sequence or miss-expression of architectural proteins, i.e., cohesin and CTCF, can lead to a wide set of human pathological phenotypes (Singh and Gerton, 2015). For example, somatic mutations in cohesin have been associated with different forms of cancer (Mullenders et al., 2015; Fisher et al., 2017; Carico et al., 2021), which is also related to the important role of cohesin in the separation of sister chromatids during cell division (Nasmyth and Haering, 2009). In addition, CTCF has been proposed as a tumor suppressor gene, since mutations in its coding sequence have also been detected in different types of cancer (Filippova et al., 2002). Interestingly, different missense mutations affected specific zinc-finger domains, leading to defects in the binding to the promoters of a subset of genes involved in the regulation of proliferation, but not to others. Therefore, both cohesin and CTCF play an essential role in gene regulation that prevents pathological situations.

CONSERVATION OF TADs ACROSS GENOME EVOLUTION

Several studies comparing chromatin structure in diverse species have reached the conclusion that TADs and their boundaries are largely conserved in animal genomes. The first HiC data comparing TADs in human and mouse cells found a general conservation of their boundaries (Dixon et al., 2012) and an in-depth analysis of conservation in mammals revealed that conserved TAD boundaries were associated with conserved CTCF sites, while divergent CTCF sites correlated with divergence of chromatin structure (Vietri Rudan et al., 2015). Indeed, CTCF binding at TAD boundaries is highly clustered and these sites are subjected to stronger selective constrains than other CTCF sites among closely related species (Kentepozidou et al., 2020; McArthur and Capra, 2021). Therefore, the strong selection against disruption of TAD boundaries in evolution is likely responsible for their enrichment in rearrangement breakpoints in vertebrates, being reshuffled as whole blocks during evolution (Krefting et al., 2018). A clear case of this conservation is the boundary splitting the RLs of the *Six* genes, which is conserved not only in vertebrates but also in echinoderms, illustrating the deep conservation of TADs involving important developmental genes (Gomez-Marin et al., 2015). Finally, studies analyzing highly conserved gene regulatory blocks, which are clusters of conserved non-coding elements around important developmental regulators, have revealed that a subset of TADs exhibit extreme non-coding conservation across metazoans (Harmston et al., 2017). Therefore, TADs involving developmental genes represent evolutionary conserved chromatin domains likely because they provide a scaffold for developmental RLs.

TAD RESHUFFLING UNDERLIES THE EMERGENCE OF EVOLUTIONARY NOVELTIES

As commented above, genomic rearrangements involving TAD restructuring and the associated alterations in gene expression usually entail deleterious effects. Nevertheless, they also provide a substrate for evolution and changes in genome structure may result in the gain of new functions that underlie the appearance of evolutionary novelties during species evolution (Maeso et al., 2017; **Figure 1B**). This has probably been the case with the emergence of limbs in vertebrates, for which the restructuring of TADs at the *HoxA* and *HoxD* loci has been essential. As can be inferred from slow evolving, non-vertebrate animals like amphioxus, the ancestral *Hox* locus was organized in a single TAD encompassing all *Hox* genes and the enhancers regulating their expression. However, at the origin of vertebrates this TAD was split in two located at either sides of the cluster and separating some genes from the rest, while leaving some others in the hinge between both TADs (**Figure 1B**; Acemel et al., 2016). This organization allowed the spatial and temporal segregation of regulatory inputs that explains the collinearity of the *HOX* genes (see above; Andrey et al., 2013; Lonfat et al., 2014) and likely enabled the plasticity in the usage of the Hox patterning system that was essential for vertebrate evolutionary novelties such as the development of paired appendages. On the other hand, the gnathostome-specific expression of *Shh* in the limbs was shown to be originated between the two whole-genome duplications by a translocation that linked the RL of that gene with *Lmbr1* (Irimia et al., 2012). Then, the ZRS enhancer could emerge in the intron 5 of *Lmbr1*, being critical for the emergence of paired and unpaired appendages (Letelier et al., 2018a).

Apart from TAD restructuring early in the evolution of vertebrates, other cases have been documented when comparing relatively close species. Thus, the regulatory cluster formed by *rac3b*, *rfng* and *sgca* genes emerged at the origin of the Ostariophysi fish superorder by a genomic rearrangement that brought in contact the RLs of *rac3b*/*rfng* and *sgca*, which are separated in Actinopterygii and tetrapods (Letelier et al., 2018b). Such rearrangement led to the formation of new regulatory contacts between *rac3b* and *rfng* promoters and the ancestral *sgca* RL, which was responsible to direct its expression to the hindbrain boundaries. These new regulatory interactions resulted in the co-option (which refers to the redeployment of pre-existing genetic or regulatory mechanisms for the acquisition of new functions or expression domains) of *rac3b* and *rfng* in the rhombomeres margins, thus promoting the formation of actomyosin cables characteristic of these structures (Letelier et al., 2018b). Moreover, it has been found that human brain tissue shows a subset of species-specific TADs compared with macaques that are associated with human-specific expression changes that are likely responsible of the higher complexity of the human brain (Luo et al., 2021). Similarly, interspecies differences in chromatin structure between human and chimpanzees are commonly associated with differences in gene expression (Eres et al., 2019). These differences between phylogenetically related species have also been discovered in two different species of *Drosophila* genus, *D. melanogaster* and *D. triauraria*. Only 25% of the TADs are orthologous between both species, and importantly these genomic rearrangements could be related with changes in gene expression (Torosin et al., 2020). These findings support the hypothesis that reorganization of genomic 3D structure may act as an important force in the rise of evolutionary novelties.

CONCLUDING REMARKS

Although the recent technological advances have allowed an increasingly detailed understanding of the 3D organization of the genome, many questions remain unanswered. How CREs operate within the context of TADs over their target genes in a mechanistic level remains incompletely understood, and to what extent the alterations of TADs lead to gene miss-expression is still an open debate. Moreover, our knowledge of how the restructuring of TADs leads to evolutionary innovations is limited to a few reported examples and the scarce availability of genome-wide chromosome conformation experiments in different species limits the comparative analyses of the 3D genome from an evolutionary perspective. Finally, the increasing applicability of the single cell technologies to chromosome conformation experiments will hopefully make possible to discern between chromatin structure in particular cell types and its association with cell type-specific gene expression during development and differentiation.

AUTHOR CONTRIBUTIONS

JS-P and JT reviewed the articles and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the European Research Council (ERC) under the European Union's Horizon 2020 and Seventh Framework Program FP7 Research and Innovation Programs (ERC-AdG-LS8-740041). JT was funded by a 2019 Leonardo Grant for Researchers and Cultural Creators, BBVA Foundation (The BBVA Foundation accepts no responsibility for the opinions, statements and contents included in the project and/or the results thereof, which are entirely the responsibility of the authors). JS-P was funded by a postdoctoral fellowship from Junta de Andalucía (DOC_00512).

ACKNOWLEDGMENTS

We apologize the authors whose work could not be discussed due to space limitations.

REFERENCES

Acemel, R. D., Maeso, I., and Gómez-Skarmeta, J. L. (2017). Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. *Wiley Interdiscip. Rev. Dev. Biol* 6:e265. doi: 10.1002/wdev.265

Acemel, R. D., Tena, J. J., Irastorza-Azcarate, I., Marlétaz, F., Gómez-Marín, C., De La Calle-Mustienes, E., et al. (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat. Genet.* 48, 336–341. doi: 10.1038/ng.3497

Akalin, A., Fredman, D., Arner, E., Dong, X., Bryne, J. C., Suzuki, H., et al. (2009). Transcriptional features of genomic regulatory blocks. *Genome Biol.* 10:R38. doi: 10.1186/gb-2009-10-4-r38

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., et al. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nat. Genet.* 52, 294–305. doi: 10.1038/s41588-019-0564-y

Alekseyenko, A. A., Walsh, E. M., Wang, X., Grayson, A. R., Hsi, P. T., Kharchenko, P. V., et al. (2015). The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains. *Genes Dev.* 29, 1507–1523. doi: 10.1101/gad.267583.115

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., et al. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* 340:1234167. doi: 10.1126/science.1234167

Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18, 285–298. doi: 10.1038/nrm.2017.7

Bell, A. C., West, A. G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387–396. doi: 10.1016/S0092-8674(00)81967-4

Bolt, C. C., and Duboule, D. (2020). The regulatory landscapes of developmental genes. *Development* 147:dev171736. doi: 10.1242/dev.171736

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 661–678. doi: 10.1038/nrg.2016.112

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171, 557–572.e24. doi: 10.1016/j.cell.2017.09.043

Busslinger, G. A., Stocsits, R. R., Van Der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., et al. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544, 503–507. doi: 10.1038/nature22063

Carico, Z. M., Stefan, H. C., Justice, M., Yimit, A., and Dowen, J. M. (2021). A cohesin cancer mutation reveals a role for the hinge domain in genome organization and gene expression. *PLoS Genet.* 17:e1009435. doi: 10.1371/journal.pgen.1009435

Chai, P., Yu, J., Jia, R., Wen, X., Ding, T., Zhang, X., et al. (2021). Generation of onco-enhancer enhances chromosomal remodeling and accelerates tumorigenesis. *Nucleic Acids Res.* 48, 12135–12150. doi: 10.1093/nar/gkaa1051

Chen, X., Ke, Y., Wu, K., Zhao, H., Sun, Y., Gao, L., et al. (2019). Key role for CTCF in establishing chromatin structure in human embryos. *Nature* 576, 306–310. doi: 10.1038/s41586-019-1812-0

Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., et al. (2017). PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* 20, 689–705.e9. doi: 10.1016/j.stem.2017.02.004

de Bruijn, S. E., Fiorentino, A., Ottaviani, D., Fanucchi, S., Melo, U. S., Corral-Serrano, J. C., et al. (2020). Structural Variants Create New Topological-Associated Domains and Ectopic Retinal Enhancer-Gene Contact in Dominant Retinitis Pigmentosa. *Am. J. Hum. Genet.* 107, 802–814. doi: 10.1016/j.ajhg.2020.09.002

de Wit, E., Vos, E. S. M., Holwerda, S. J. B., Valdes-Quezada, C., Verstegen, M. J. A. M., Teunissen, H., et al. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* 60, 676–684. doi: 10.1016/j.molcel.2015.09.023

Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., Jahn, K. S., et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* 158, 849–860. doi: 10.1016/j.cell.2014.05.050

Despang, A., Schöpflin, R., Franke, M., Ali, S., Jerković, I., Paliou, C., et al. (2019). Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* 51, 1263–1271. doi: 10.1038/s41588-019-0466-z

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336. doi: 10.1038/nature14222

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380. doi: 10.1038/nature11082

Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159, 374–387. doi: 10.1016/j.cell.2014.09.030

Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., et al. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 547, 232–235. doi: 10.1038/nature23263

Eagen, K. P., and French, C. A. (2021). Supercharging BRD4 with NUT in carcinoma. *Oncogene* 40, 1396–1408. doi: 10.1038/s41388-020-01625-0

Eres, I. E., Luo, K., Hsiao, C. J., Blake, L. E., and Gilad, Y. (2019). Reorganization of 3D genome structure may contribute to gene regulatory evolution in primates. *PLoS Genet.* 15:e1008278. doi: 10.1371/journal.pgen.1008278

Espinola, S. M., Götz, M., Bellec, M., Messina, O., Fiche, J. B., Houbron, C., et al. (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. *Nat. Genet.* 53, 477–486. doi: 10.1038/s41588-021-00816-z

Filippova, G. N., Fagerlie, S., Klenova, E. M., Myers, C., Dehner, Y., Goodwin, G., et al. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell. Biol.* 16, 2802–2813. doi: 10.1128/mcb.16.6.2802

Filippova, G. N., Ulmer, J. E., Moore, J. M., Ward, M. D., Hu, Y. J., Neiman, P. E., et al. (2002). Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res.* 62, 48–52.

Fisher, J. B., McNulty, M., Burke, M. J., Crispino, J. D., and Rao, S. (2017). Cohesin Mutations in Myeloid Malignancies. *Trends Cancer* 3, 282–293. doi: 10.1016/j.trecan.2017.02.006

Franke, M., De la Calle-Mustienes, E., Neto, A., Acemel, R. D., Santos-Pereira, J. M., and Gómez-Skarmeta, J. L. (2020). CTCF knockout in zebrafish induces alterations in 1 regulatory landscapes and developmental gene expression 2 3. *bioRxiv* [Preprint]. doi: 10.1101/2020.09.08.282707

Franke, M., and Gomez-Skarmeta, J. L. (2018). An evolutionary perspective of regulatory landscape dynamics in development and disease. *Curr. Opin. Cell Biol.* 55, 24–29. doi: 10.1016/j.ceb.2018.06.009

Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schopflin, R., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* 538, 265–269. doi: 10.1038/nature19800

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15, 2038–2049. doi: 10.1016/j.celrep.2016.04.085

Furlong, E. E. M., and Levine, M. (2018). Developmental enhancers and chromosome topology. *Science* 361, 1341–1345. doi: 10.1126/science.aau0320

Galupa, R., and Heard, E. (2017). Topologically Associating Domains in Chromosome Architecture and Gene Regulatory Landscapes during Development, Disease, and Evolution. *Cold Spring Harb. Symp. Quant. Biol.* 82, 267–278. doi: 10.1101/sqb.2017.82.035030

Gassler, J., Brandão, H. B., Imakaev, M., Flyamer, I. M., Ladstätter, S., Bickmore, W. A., et al. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* 36, 3600–3618. doi: 10.15252/embj.201798083

Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R. R., Korbel, J. O., and Furlong, E. E. M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* 51, 1272–1282. doi: 10.1038/s41588-019-0462-3

Ghavi-Helm, Y., Klein, F. A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., et al. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512, 96–100. doi: 10.1038/nature13417

Gomez-Marin, C., Tena, J. J., Acemel, R. D., Lopez-Mayorga, M., Naranjo, S., de la Calle-Mustienes, E., et al. (2015). Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7542–7547. doi: 10.1073/pnas.1505463112

Haarhuis, J. H. I., van der Weide, R. H., Blomen, V. A., Yáñez-Cuna, J. O., Amendola, M., van Ruiten, M. S., et al. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* 169, 693–707.e14. doi: 10.1016/j.cell.2017.04.013

Haberle, V., Arnold, C. D., Pagani, M., Rath, M., Schernhuber, K., and Stark, A. (2019). Transcriptional cofactors display specificity for distinct types of core promoters. *Nature* 570, 122–126. doi: 10.1038/s41586-019-1210-7

Hadjur, S., Williams, L. M., Ryan, N. K., Cobb, B. S., Sexton, T., Fraser, P., et al. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460, 410–413. doi: 10.1038/nature08079

Harmston, N., Ing-Simmons, E., Tan, G., Perry, M., Merkenschlager, M., and Lenhard, B. (2017). Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation. *Nat. Commun.* 8:441. doi: 10.1038/s41467-017-00524-5

Hnisz, D., Weintraub, A. S., Day, D. S., Valton, A. L., Bak, R. O., Li, C. H., et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351, 1454–1458. doi: 10.1126/science.aad9024

Hsieh, T. H. S., Cattoglio, C., Slobodyanyuk, E., Hansen, A. S., Rando, O. J., Tjian, R., et al. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol. Cell* 78, 539–553.e8. doi: 10.1016/j.molcel.2020.03.002

Hug, C. B., Grimaldi, A. G., Kruse, K., and Vaquerizas, J. M. (2017). Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* 169, 216–228.e19. doi: 10.1016/j.cell.2017.03.024

Ibn-Salem, J., Köhler, S., Love, M. I., Chung, H. R., Huang, N., Hurles, M. E., et al. (2014). Deletions of chromosomal regulatory boundaries are associated with congenital disease. *Genome Biol.* 15:423. doi: 10.1186/s13059-014-0423-1

Ibrahim, D. M., and Mundlos, S. (2020). Three-dimensional chromatin in disease: what holds us together and what drives us apart? *Curr. Opin. Cell Biol.* 64, 1–9. doi: 10.1016/j.ceb.2020.01.003

Ing-Simmons, E., Vaid, R., Bing, X. Y., Levine, M., Mannervik, M., and Vaquerizas, J. M. (2021). Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning. *Nat. Genet.* 53, 487–499. doi: 10.1038/s41588-021-00799-x

Irimia, M., Royo, J. L., Burguera, D., Maeso, I., Gomez-Skarmeta, J. L., and Garcia-Fernandez, J. (2012). Comparative genomics of the Hedgehog loci in chordates and the origins of Shh regulatory novelties. *Sci. Rep.* 2:433. doi: 10.1038/srep00433

Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M., and Epstein, D. J. (2006). A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* 133, 761–772. doi: 10.1242/dev.02239

Jerkovic', I., and Cavalli, G. (2021). Understanding 3D genome organization by multidisciplinary methods. *Nat. Rev. Mol. Cell Biol.* doi: 10.1038/s41580-021-00362-w [Epub ahead of print].

Ju, L., Wing, J., Taylor, E., Brandt, R., Slijepcevic, P., Horsch, M., et al. (2013). SMC6 is an essential gene in mice, but a hypomorphic mutant in the ATPase domain has a mild phenotype with a range of subtle abnormalities. *DNA Repair* 12, 356–366. doi: 10.1016/j.dnarep.2013.02.006

Kaaij, L. J. T., van der Weide, R. H., Ketting, R. F., and de Wit, E. (2018). Systemic Loss and Gain of Chromatin Architecture throughout Zebrafish Development. *Cell Rep.* 24, 1–10.e4. doi: 10.1016/j.celrep.2018.06.003

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435. doi: 10.1038/nature09380

Katainen, R., Dave, K., Pitkänen, E., Palin, K., Kivioja, T., Välimäki, N., et al. (2015). CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* 47, 818–821. doi: 10.1038/ng.3335

Kentepozidou, E., Aitken, S. J., Feig, C., Stefflova, K., Ibarra-Soria, X., Odom, D. T., et al. (2020). Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. *Genome Biol.* 21:5. doi: 10.1186/s13059-019-1894-x

Krefting, J., Andrade-Navarro, M. A., and Ibn-Salem, J. (2018). Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biol.* 16:87. doi: 10.1186/s12915-018-0556-x

Krietenstein, N., Abraham, S., Venev, S. V., Abdennur, N., Gibcus, J., Hsieh, T. H. S., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. *Mol. Cell* 78, 554–565.e7. doi: 10.1016/j.molcel.2020.03.003

Kubo, N., Ishii, H., Xiong, X., Bianco, S., Meitinger, F., Hu, R., et al. (2021). Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat. Struct. Mol. Biol.* 28, 152–161. doi: 10.1038/s41594-020-00539-5

Laugsch, M., Bartusel, M., Rehimi, R., Alirzayeva, H., Karaolidou, A., Crispatzu, G., et al. (2019). Modeling the Pathological Long-Range Regulatory Effects of Human Structural Variation with Patient-Specific hiPSCs. *Cell Stem Cell* 24, 736–752.e12. doi: 10.1016/j.stem.2019.03.004

Letelier, J., de la Calle-Mustienes, E., Pieretti, J., Naranjo, S., Maeso, I., Nakamura, T., et al. (2018a). A conserved Shh cis-regulatory module highlights a common developmental origin of unpaired and paired fins. *Nat. Genet.* 50, 504–509. doi: 10.1038/s41588-018-0080-5

Letelier, J., Terriente, J., Belzunce, I., Voltes, A., Undurraga, C. A., Polvillo, R., et al. (2018b). Evolutionary emergence of the rac3b/rfng/sgca regulatory cluster refined mechanisms for hindbrain boundaries formation. *Proc. Natl. Acad. Sci. U. S. A.* 115, E3731–E3740. doi: 10.1073/pnas.1719885115

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., et al. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* 12, 1725–1735. doi: 10.1093/hmg/ddg180

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. doi: 10.1126/science.1181369

Lonfat, N., and Duboule, D. (2015). Structure, function and evolution of topologically associating domains (TADs) at HOX loci. *FEBS Lett.* 589, 2869–2876. doi: 10.1016/j.febslet.2015.04.024

Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* 346, 1004–1006. doi: 10.1126/science.1257493

Luo, X., Liu, Y., Dang, D., Hu, T., Hou, Y., Meng, X., et al. (2021). 3D Genome of macaque fetal brain reveals evolutionary innovations during primate corticogenesis. *Cell* 184, 723–740.e21. doi: 10.1016/j.cell.2021.01.001

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–1025. doi: 10.1016/j.cell.2015.04.004

Maeso, I., Acemel, R. D., and Gomez-Skarmeta, J. L. (2017). Cis-regulatory landscapes in development and evolution. *Curr. Opin. Genet. Dev.* 43, 17–22. doi: 10.1016/j.gde.2016.10.004

McArthur, E., and Capra, J. A. (2021). Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. *Am. J. Hum. Genet.* 108, 269–283. doi: 10.1016/j.ajhg.2021.01.001

McCord, R. P., Kaplan, N., and Giorgetti, L. (2020). Chromosome Conformation Capture and Beyond: toward an Integrative View of Chromosome Structure and Function. *Mol. Cell* 77, 688–708. doi: 10.1016/j.molcel.2019.12.021

Merkenschlager, M., and Nora, E. P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu. Rev. Genomics Hum. Genet.* 17, 17–43. doi: 10.1146/annurev-genom-083115-022339

Moore, J. M., Rabaia, N. A., Smith, L. E., Fagerlie, S., Gurley, K., Loukinov, D., et al. (2012). Loss of maternal CTCF is associated with peri-implantation lethality of Ctcf null embryos. *PLoS One* 7:e34915. doi: 10.1371/journal.pone.0034915

Mullenders, J., Aranda-Orgilles, B., Lhoumaud, P., Keller, M., Pae, J., Wang, K., et al. (2015). Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J. Exp. Med.* 212, 1833–1850. doi: 10.1084/jem.20151323

Nakamura, R., Motai, Y., Kumagai, M., Nishiyama, H., Durand, N. C., Kondo, K., et al. (2018). CTCF looping is established during gastrulation in medaka embryos. *bioRxiv* [preprint]. doi: 10.1101/454082

Nasmyth, K., and Haering, C. H. (2009). Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 43, 525–558. doi: 10.1146/annurev-genet-102108-134233

Nora, E. P., Caccianini, L., Fudenberg, G., So, K., Kameswaran, V., Nagle, A., et al. (2020). Molecular basis of CTCF binding polarity in genome folding. *Nat. Commun.* 11:5612. doi: 10.1038/s41467-020-19283-x

Nora, E. P., Goloborodko, A., Valton, A. L., Gibcus, J. H., Uebersohn, A., Abdennur, N., et al. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 169, 930–944.e22. doi: 10.1016/j.cell.2017.05.004

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385. doi: 10.1038/nature11049

Oudelaar, A. M., Beagrie, R. A., Gosden, M., de Ornellas, S., Georgiades, E., Kerry, J., et al. (2020). Dynamics of the 4D genome during in vivo lineage specification and differentiation. *Nat. Commun.* 11:2722. doi: 10.1038/s41467-020-16598-7

Pachano, T., Sánchez-Gaya, V., Mariner-Faulí, M., Ealo, T., Asenjo, H. G., Respuela, P., et al. (2020). Orphan CpG islands boost the regulatory activity of poised enhancers and dictate the responsiveness of their target genes. *bioRxiv* [preprint]. doi: 10.1101/2020.08.05.237768

Paliou, C., Guckelberger, P., Schöpflin, R., Heinrich, V., Esposito, A., Chiariello, A. M., et al. (2019). Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl. Acad. Sci. U. S. A.* 116, 12390–12399. doi: 10.1073/pnas.1900672116

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132, 422–433. doi: 10.1016/j.cell.2008.01.011

Phillips, J. E., and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* 137, 1194–1211. doi: 10.1016/j.cell.2009.06.001

Rada-Iglesias, A., Grosveld, F. G., and Papantonis, A. (2018). Forces driving the three-dimensional folding of eukaryotic genomes. *Mol. Syst. Biol.* 14:e8214. doi: 10.15252/msb.20188214

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665–1680. doi: 10.1016/j.cell.2014.11.021

Rao, S. S. P., Huang, S. C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K. R., et al. (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* 171, 305–320.e24. doi: 10.1016/j.cell.2017.09.026

Rosencrance, C. D., Ammouri, H. N., Yu, Q., Ge, T., Rendleman, E. J., Marshall, S. A., et al. (2020). Chromatin Hyperacetylation Impacts Chromosome Folding by Forming a Nuclear Subcompartment. *Mol. Cell* 78, 112–126.e12. doi: 10.1016/j.molcel.2020.03.018

Rowley, M. J., and Corces, V. G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19, 789–800. doi: 10.1038/s41576-018-0060-8

Sagai, T., Amano, T., Tamura, M., Mizushina, Y., Sumiyama, K., and Shiroishi, T. (2009). A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* 136, 1665–1674. doi: 10.1242/dev.032714

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132, 797–803. doi: 10.1242/dev.01613

Sanborn, A. L., Rao, S. S. P., Huang, S. C., Durand, N. C., Huntley, M. H., Jewett, A. I., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.* 112, E6456–E6465. doi: 10.1073/pnas.1518552112

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51–56. doi: 10.1038/nature24281

Serebreni, L., and Stark, A. (2021). Insights into gene regulation: from regulatory genomic elements to DNA-protein and protein-protein interactions. *Curr. Opin. Cell Biol.* 70, 58–66. doi: 10.1016/j.ceb.2020.11.009

Singh, V. P., and Gerton, J. L. (2015). Cohesin and human disease: lessons from mouse models. *Curr. Opin. Cell Biol.* 37, 9–17. doi: 10.1016/j.ceb.2015.08.003

Spielmann, M., Lupiáñez, D. G., and Mundlos, S. (2018). Structural variation in the 3D genome. *Nat. Rev. Genet.* 19, 453–467. doi: 10.1038/s41576-018-0007-0

Spitz, F., Gonzalez, F., and Duboule, D. (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* 113, 405–417. doi: 10.1016/s0092-8674(03)00310-6

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., et al. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* 24, 390–400. doi: 10.1101/gr.163519.113

Torosin, N. S., Anand, A., Golla, T. R., Cao, W., and Ellison, C. E. (2020). 3D genome evolution and reorganization in the Drosophila melanogaster species group. *PLoS Genet.* 16:e1009229. doi: 10.1371/journal.pgen.1009229

van der Weide, R. H., and de Wit, E. (2019). Developing landscapes: genome architecture during early embryogenesis. *Curr. Opin. Genet. Dev.* 55, 39–45. doi: 10.1016/j.gde.2019.04.009

Vicente-Garcia, C., Villarejo-Balcells, B., Irastorza-Azcarate, I., Naranjo, S., Acemel, R. D., Tena, J. J., et al. (2017). Regulatory landscape fusion in rhabdomyosarcoma through interactions between the PAX3 promoter and FOXO1 regulatory elements. *Genome Biol.* 18:106. doi: 10.1186/s13059-017-1225-z

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D. T., Tanay, A., et al. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10, 1297–1309. doi: 10.1016/j.celrep.2015.02.004

Weintraub, A. S., Li, C. H., Zamudio, A. V., Sigova, A. A., Hannett, N. M., Day, D. S., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* 171, 1573–1588.e28. doi: 10.1016/j.cell.2017.11.008

Wike, C. L., Guo, Y., Tan, M., Nakamura, R., Shaw, D. K., Díaz, N., et al. (2021). Chromatin architecture transitions from zebrafish sperm through early embryogenesis. *Genome Res.* 31, 981–994. doi: 10.1101/gr.269860.120

Williamson, I., Kane, L., Devenney, P. S., Flyamer, I. M., Anderson, E., Kilanowski, F., et al. (2019). Developmentally regulated Shh expression is robust to TAD perturbations. *Development* 146:dev179523. doi: 10.1242/dev.179523

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D. A., Stocsits, R. R., Tang, W., et al. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* 36, 3573–3599. doi: 10.15252/embj.201798004

Zhan, Y., Mariani, L., Barozzi, I., Schulz, E. G., Blüthgen, N., Stadler, M., et al. (2017). Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res.* 27, 479–490. doi: 10.1101/gr.212803.116

Conflict of Interest: 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 © 2021 Tena and Santos-Pereira. 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) and the copyright owner(s) 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 Cell and Developmental Biology

REVIEW
published: 06 July 2021
doi: 10.3389/fcell.2021.713918

New Genes Born-In or Invading Vertebrate Genomes

*Carlos Herrera-Úbeda and Jordi Garcia-Fernàndez**

Department of Genetics, Microbiology and Statistics, Faculty of Biology, and Institute of Biomedicine (IBUB), University of Barcelona, Barcelona, Spain

Which is the origin of genes is a fundamental question in Biology, indeed a question older than the discovery of genes itself. For more than a century, it was uneven to think in origins other than duplication and divergence from a previous gene. In recent years, however, the intersection of genetics, embryonic development, and bioinformatics, has brought to light that *de novo* generation from non-genic DNA, horizontal gene transfer and, noticeably, virus and transposon invasions, have shaped current genomes, by integrating those newcomers into old gene networks, helping to shape morphological and physiological innovations. We here summarized some of the recent research in the field, mostly in the vertebrate lineage with a focus on protein-coding novelties, showing that the placenta, the adaptative immune system, or the highly developed neocortex, among other innovations, are linked to *de novo* gene creation or domestication of virus and transposons. We provocatively suggest that the high tolerance to virus infections by bats may also be related to previous virus and transposon invasions in the bat lineage.

Keywords: transposon-domestication, virus-invasion, Bex-Tceal, gene_birth, junk_DNA

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
University of Málaga, Spain

Reviewed by:
Bastian Fromm,
Stockholm University, Sweden
Daqi Yu,
Institute of Zoology, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Jordi Garcia-Fernàndez
jordigarcia@ub.edu

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *24 May 2021*
Accepted: *15 June 2021*
Published: *06 July 2021*

Citation:
Herrera-Úbeda C and Garcia-Fernàndez J (2021) New Genes Born-In or Invading Vertebrate Genomes. Front. Cell Dev. Biol. 9:713918. doi: 10.3389/fcell.2021.713918

INTRODUCTION: NEARLY TWO CENTURIES OF HONEYMOONS AND DIVORCES OF SCIENTIFIC FIELDS

That Charles Darwin changed the world with his book on the origin of the species (Darwin, 1859), and quaked human thinking with its “Descent of man” (Darwin, 1871) is beyond doubt. However, he did not know what a gene was. Now, we know what a gene is, but we do not really understand where genes come from. What Darwin was aware, however, was of the relevance of Embryology and, in fact, he based his evolutionary theory in part on the old observations of embryos similarities, regardless of adult similarities, of Karl Enst von Baer (1828). In Darwin words: “community of embryonic structure reveals community of descent,” (Darwin, 1859). As a landmark example that validated the evolution/development marriage was the discovery by the Russian embryologist Kowalewski of a notochord in the urochordates (Kowalewski, 1866), noticeably even to non-specialist readers: “One could hardly open a scientific Journal or any popular essay on Natural History without meeting some allusion to the Ascidians as our ancestors” (Agassiz, 1874). Also, Haeckel and his biogenetic law (Haeckel, 1866) and his famous statement “ontogeny recapitulates phylogeny” became very popular. Therefore, at the end of the XIX century, the evolution of species was linked to developmental biology and developmental biology to cell behavior but, however, nothing about the evolution of genes, the ultimate actors of living beings.

The beginning of the XX century was dramatic for Evolution. Genes were (re)discovered, but the nascent new science, genetics, was a very disturbing third party to the evolution/development marriage and thus for comparative biology: genes were believed to have nothing to do with embryonic development (Bateson, 1922). It is true that mutation rates, allele frequencies, vast amounts of mathematics, and exquisite research in wild animal and plant population insufflate and developed a well-established evolutionary thinking, named Evolutionary Synthesis or Modern Synthesis (Huxley, 1942). Later, several Extended Synthesis that included a variety of new concepts in addition to classic Mendelian Genetics were added (Eldredge and Gould, 1972; Mayr and Provine, 1980). However, where new genes came from was not a central issue for the new or extended synthesis. Still, after Watson and Crick (1953), it was believed that genes from flies and mouse, as an example, had to be extremely different and mainly unrelated. In addition, it was thought that the human genome had to include many more genes than the fly genome, as it was obvious in terms of body plan complexity (from a human point of view).

This "old" view of evolution was shaken in 1978 by a major article that merited a Nobel prize 20 years later, the discovery of the Hox cluster (Lewis, 1978). In 1984 the homeobox was identified molecularly (McGinnis et al., 1984; Scott and Weiner, 1984) and Hox genes from mouse and human, intriguingly similar to the fly genes, discovered. The homeobox, named the "Rosetta Stone" of Biology (Garcia-Fernàndez, 2005) was the launching point of a new discipline and, indeed, the start of the reconciliation of Evolution and Embryology. With Genetics as the major actor of the so-called New-new Synthesis, New extended synthesis or Evo-Devo, the rationale was simple: if developmental genes regulate development, development regulates morphology and physiology and evolution relies on morphology and physiology, then understanding the evolution of developmental genes is the crux to understanding evolution (Baguñà and Garcia-Fernàndez, 2003). Interestingly indeed, the central period of embryonic development, the so-called phylotypic stage, was the stage where crucial (and highly conserved) developmental genes were expressed, suggesting that only earlier and final periods of development were highly evolvable. This was named the hourglass model that related development and evolution (Raff, 1996; Wagner, 2014) and the extended Hox family that is expressed there as the "zootype" (Slack et al., 1993).

In this new field, myriads of reports highlighted the conservation of genes, gene networks among large evolutionary distances and overall number of protein-coding genes in metazoans, creating a paradox at the beginning of the XX century, especially in the animal kingdom: if animals share all this, why are they so different? The solution to the paradox may well rely upon changes of gene regulation in a broad sense, from changes in *cis*-regulation, post-transcriptional regulation, or effects of gene duplication and an ever-growing list of new regulatory elements, with microRNAs, genome topology, epigenomics, long non-coding RNAs or gene editing as the latest players (Heimberg et al., 2008; Wang and Chang, 2011; Irimia et al., 2012; Cavalli and Misteli, 2013; Fromm et al., 2015; Liscovitch-Brauer et al., 2017; Deline et al., 2018). However, not much attention was given to a kind of change in genome functioning which may have an unforeseeable impact in evolution: the birth of new genes.

For half a century, most scientists believed that new protein-coding genes arise as a result of mutations in existing protein-coding genes. It was considered nearly impossible for anything as complex as a functional new protein to arise from scratch (Jacob, 1977). This was puzzling to the understanding of the evolution of metazoans, especially for those major morphological transitions that constitute evolutionary key points bound to radical innovations (Simakov and Kawashima, 2017), from the origins of multicellularity (Stanley, 1973) to the mammal placenta (Roberts et al., 2016). Fortunately, and thanks to the present affordability of new high throughput technologies, an explosion of transcriptomic and genomic data from key phylogenetic species have allowed us to widen our understanding of how such novelties have arisen (Van Oss and Carvunis, 2019).

These disruptive novelties are often accompanied by the apparition of new genes that integrate into the current gene networks (Lavialle et al., 2013; Parrish and Tomonaga, 2018; Zhang et al., 2019; Navas-Pérez et al., 2020; Pascual-Carreras et al., 2020). These new genes (**Figure 1**) can of course be paralogs caused by duplication of genes already present in the genome which could allow the duplicated gene to change its sequence or its regulation (Ohno, 1970; Putnam et al., 2008; Jimenez-Delgado et al., 2009). But they can be fully *de novo* genes, which have no evident homology with any other gene within the studied species or their close relatives: the also known as orphan or taxonomically restricted genes (Singh and Syrkin Wurtele, 2020). We classify here those protein-coding new players found in old genomes from its origin and summarized the knowledge of their impact, with a special focus on its involvement in vertebrate evolution, highlighting the role of virus, virus-related elements, and transposons, when invading vertebrate and mammalian old genomes.

TWO SIDES OF THE SAME COIN: NEW GENES FROM SCRATCH AND NEW GENES FROM ALREADY EXISTING GENIC REGIONS

Although taxonomically restricted genes are a very heterogeneous group, they can be classified according to their origin (**Figure 2**). The true *de novo* genes emerged from ancestrally non-genic regions (wrongly called "junk DNA" in the past). These new genes can arise by mechanisms still poorly understood that involve a genomic region gaining both, transcriptional activity (maybe by transcriptional leaking) and an ORF (open reading frame) in either order (although the ORF is not necessary for non-coding genes). That would seem to be the case of the gene *blitzschnell* found in planarians (Pascual-Carreras et al., 2020) although with the current information we can not infer if the ORF originated first, or instead it was the transcriptional activity. Another mechanism that could incorporate non-genic material to the gene repertoire would be overprinting. In this process, a new ORF is created overlapping an existing one but in a different frame, resulting in two or more genes with overlapping coordinates but with substantially different amino acid sequences (Delaye et al., 2008). Finally, non-genic material can be added to an already existing gene through exonization, in which new exons are generated by random

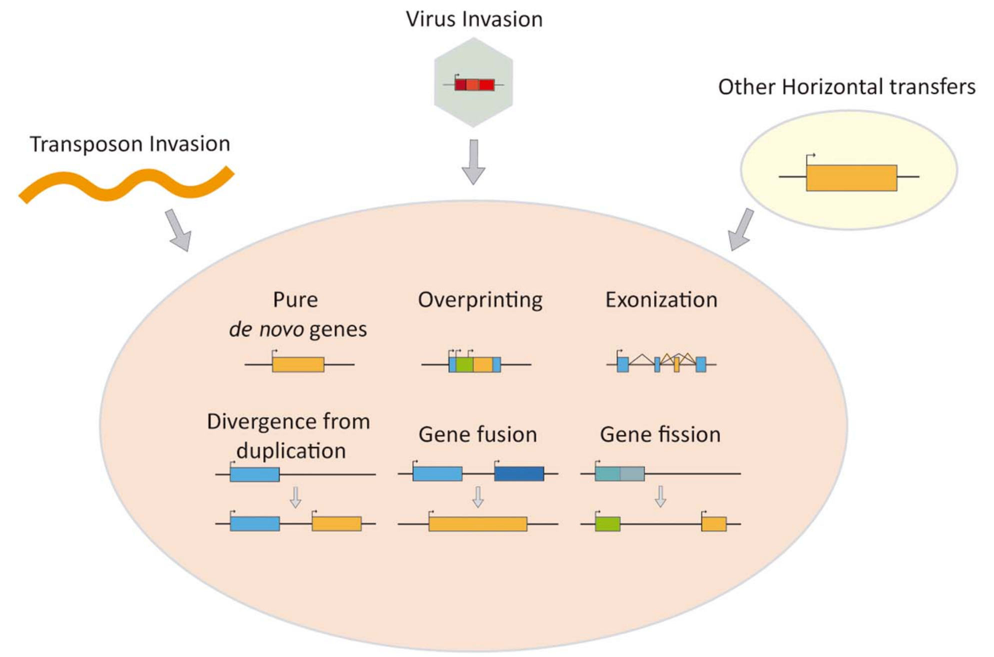


FIGURE 1 | Schematics of mechanisms for generating *de novo* genes.

mutations in non-genic DNA, like most alternative exons regulated by NOVA (Irimia et al., 2011), a splicing factor known to be responsible for the inclusion of previously non-extant exons using cryptic splice sites near to the Nova-binding motifs. Or the well know example of the extra exon of TRPV1 (similar to the one present in Laurasiatheria mammals) that has been co-opted in vampire bats to detect hot spots on warm-blooded prey (Gracheva et al., 2011).

Meanwhile, other taxonomically restricted genes can emerge from several other mechanisms (**Figure 2**), such as extreme divergence from a previous duplication, gene fusion, gene fission, horizontal gene transfer, retroposition or domestication of viral or transposable elements. Remarkably, the process of domestication has been reported to be responsible for disruptive innovations, as independently evolving genes integrate into a new gene framework.

DOMESTICATED RETROVIRAL PROTEINS THROUGH EVOLUTION

Endogenous viral elements are remnants from the integration of retroviruses into the genome and are quite abundant in the animal genomes (Katzourakis and Gifford, 2010). For example, the human genome is formed approximately by 8% of endogenous retroviral sequences (Lander et al., 2001). Throughout evolution, most of the genes from these sequences lose their function, but some of them are captured and "domesticated" in a process called exaptation. Among the examples of domesticated viral genes, the group that has most clearly influenced the evolution of mammals are syncytins (Lavialle et al., 2013). These captured viral proteins are the product of an envelope gene of retroviruses ancestrally endogenized. The envelope glycoprotein (Env) is crucial in the process of viral entry in enveloped viruses and induces fusion of the virion envelope with the cell plasmatic membrane (Sha et al., 2000). Within the human genome several Env genes can be found, but only two of them have a placental-specific expression and induce the formation of syncytia (**Figure 3**). Being the placenta such a defining organ in placental mammals, the syncytins responsible for its development could be expected to be orthologs in the different species, but that is not the case (Lavialle et al., 2013). Primate and mouse syncytins are not syntenic, and there is evidence pointing to independent capture events in the ancestors of each clade, as well as in the Scincidae genus Mabuya (Cornelis et al., 2017). In fact, in mammals, the different capture events can be linked with the four different main types of placental structures. In the same way, the differences between the lizard placenta present in the genus Mabuya and the mammal placenta can be traced to a completely different capture of Env genes (Cornelis et al., 2017).

Similarly, Arc genes mediate intercellular communication and synaptic plasticity via extracellular vesicles (Parrish and Tomonaga, 2018), and are homologous to the Group-specific antigen (Gag) polyproteins. In retroviruses, capsids are necessary for cellular infection and their assembly is mainly mediated by Gag (Rose et al., 2020). The similarities between Arc and Gag are not restricted to sequence, as Arc is able to spontaneously assemble into a structure that resembles to a capsid (Pastuzyn et al., 2018). In fact, Arc not only forms these capsid-like structures but also encapsulates any mRNA present during their formation. These processes allow the traffic of RNA molecules

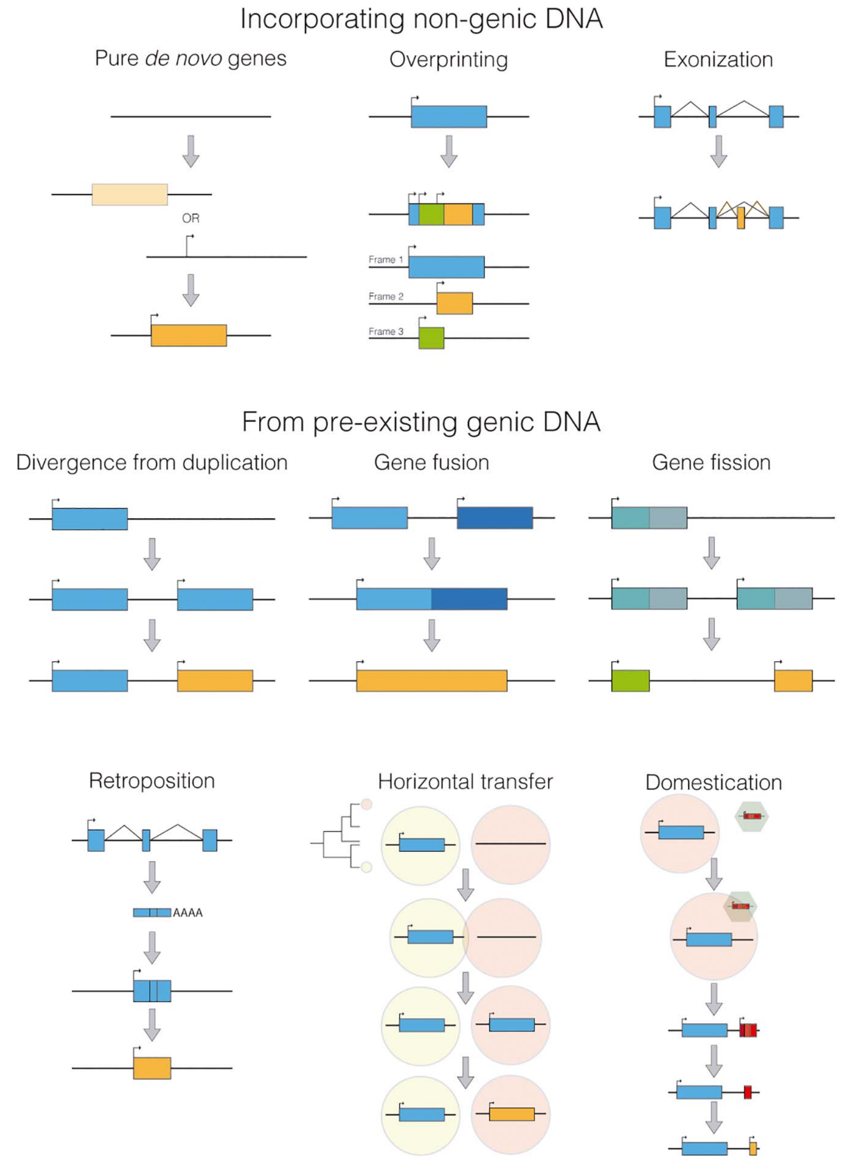

Incorporating non-genic DNA
Pure *de novo* genes
Overprinting
Exonization
OR
Frame 1
Frame 2
Frame 3
From pre-existing genic DNA
Divergence from duplication
Gene fusion
Gene fission
Retroposition
AAAA
Horizontal transfer
Domestication

FIGURE 2 | Different mechanisms for generating taxonomically restricted genes.

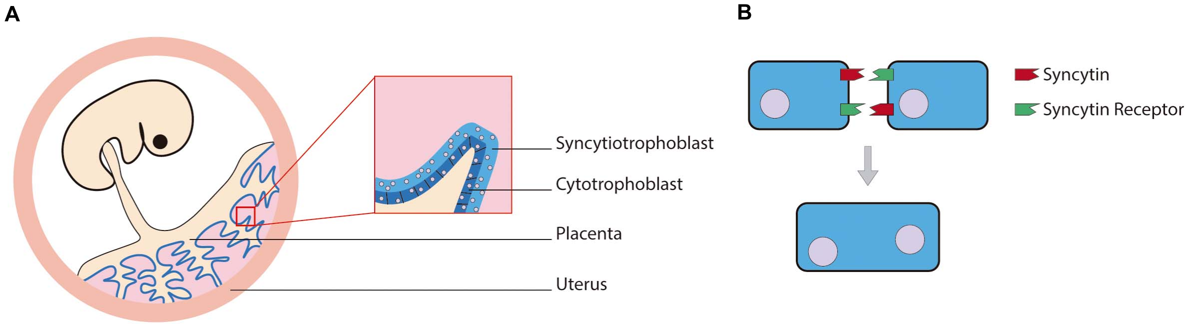

A
Syncytiotrophoblast
Cytotrophoblast
Placenta
Uterus
B
Syncytin
Syncytin Receptor

FIGURE 3 | (A) Human example of the syncytiotrophoblast location in the placenta alongside **(B)** a diagram of the syncytium formation mediated by syncytins.

between nervous system cells (Ashley et al., 2018). Regarding the capture and evolution of this viral protein, phylogenetic analyses showed at least two independent capture events that took place in the ancestors of tetrapods and in that of schizophorans (Pastuzyn et al., 2018). In both cases, the co-option of Arc led to similar functions of RNA trafficking in the nervous system. In both lineages the closest hit was a Ty3/gypsy retrotransposon, but tetrapod Arc grouped with the fish Ty3/gypsy, while fly Arc grouped with the insect Ty3/Gypsy, indicating that while sharing significant homology with the Gag protein, it seems to have originated from the Ty3/gypsy retrotransposons from each lineage.

TRANSPOSON-DERIVED NOVELTIES IN VERTEBRATE EVOLUTION

As with the endogenous viral elements, transposable elements can be a source of disruptive innovations. The process of transposition can place sequences near new promoters or generate new fusion proteins. In the evolution of the adaptative immune system of vertebrates, the domestication of the RAG (recombination-activating genes) transposon was instrumental for the V(D)J recombination system, which is a process that makes possible the diversity of antibodies and T cell receptors present in the vertebrate adaptative immune system (Zhang et al., 2019). The current model, supported by the presence of ProtoRAG in the pre-vertebrate amphioxus, is that an ancestral Transib transposon with RAG1-like ORF and terminal inverted repeats, similar to the recombination signal sequences present in V(D)J, captured a RAG2-like ORF to form the ancestral RAG transposon. This event, which took place in an early deuterostome (Carmona and Schatz, 2017), was followed in jawed vertebrates by the insertion of the RAG transposon into a gene encoding an immunoglobulin-domain receptor, among other changes that suppressed RAG transposition activity to finally constitute the V(D)J recombination system.

Another example of a domesticated transposable element is the paired box (PAX) family (Paixão-Côrtes et al., 2015). These homeotic genes were discovered in 1986 (Bopp et al., 1986) and have been proved to be master regulators of the development in metazoans (Dahl et al., 1997). There are several PAX genes, and they were thought to have a monophyletic origin which was initially set at the beginning of metazoan diversification (Hoshiyama et al., 2007). However, PAX-like genes were found in protozoans placing the ancestral PAX down to the pre-metazoan era (Wang et al., 2010). The origin of the ancestral PAX gene is characterized by the domestication of a Tc1/mariner transposon, an ancient and widespread transposon family present in metazoans as well as in plants and protozoans (Garcia-Fernàndez et al., 1993). The Tc1/mariner transposase is similar to the PAX DNA-binding paired domain and its capture was probably posterior to the formation of the other two characteristics domains.

The Pax family represents not the only transcription factor formed by fusion of a transposase domain with another gene. The host-transposase fusion (HTF) genes are a group of genes that arose most probably from exon shuffling (Cosby et al., 2021), where a transposable element landed within an intron of an existing gene. Once there, the splicing machinery used the splice acceptor/donor sites pre-existing in the integrated transposon. The resemblances in the origin of several HTFs suggest then that DNA transposons are prone to be captured via alternative splicing. In fact, in tetrapod lineages, 106 distinct HTFs have been identified recently from 106 independent events (Cosby et al., 2021).

WHEN CAN NEW GENES INTEGRATE INTO EXISTING NETWORKS?

As we have seen, these kinds of newborn genes are not as rare as initially thought and can shape the evolution of entire groups of animals. But in order to generate a disruptive novelty, a new gene not only has to be born but to be integrated into an already existing gene network. Genes mainly involved in the phylotypic stage, the stage of development shared by all members of the phyla (Duboule, 1994) (the elongated neurula stage in the case of vertebrates) are ancient genes forming highly conserved gene networks (Irie and Sehara-Fujisawa, 2007), where the slightest variation could wreak havoc in the most crucial stages of development. Oppositely, gene networks acting in very early or very late stages of development tend to show more variation and have less conserved or even new genes implicated (Irie and Kuratani, 2011). In this context, the "hourglass model" (Drost et al., 2017), recently also found at the gene regulatory level (Liu et al., 2020), suggest that if a new gene is born, it may preferentially end up functioning at very early (e.g., placenta) or very late stages of the life cycle (e.g., adaptative immunity, synaptic plasticity, body size regulation).

A good example besides the ones already mentioned would be the aforementioned taxonomically restricted gene family, blitzschnell (Pascual-Carreras et al., 2020). Found only within the order Tricladida (planarians), this family is composed of 11 genes and four pseudogenes. It can be further divided into five subfamilies, with one of them organized in a cluster formed by tandem duplication events. Three of the subfamilies are coding and have been reported to regulate the growth/degrowth according to nutrient intake. Thus, these *de novo* genes have been integrated into an evolutionary conserved metabolic network, the insulin/Akt/mTOR network responsible for growth (Saxton and Sabatini, 2017) and other mechanisms at a late stage of the life cycle.

A WHOLE NEW CLUSTER FROM A DOMESTICATED TRANSPOSABLE ELEMENT: IMPACT INTO THE EVOLUTION OF THE EUTHERIAN BRAIN

Recently Navas-Pérez et al. (2020) presented evidence of the expansion of a domestication event into the Bex/Tceal multigenic family, constituting a cluster of 14 genes in the X chromosome of the ancestor of eutherians, after the divergence of the marsupial-placental clades. The domestication event proposed (**Figure 4**) consists of six steps: (i) A proto-BGW motif (Winter and Ponting, 2005) existed upstream of the alpha-galactosidase (Gla) promoter (P α) in the X chromosome of the ancestor of eutherians and metatherians; (ii) In the eutherian lineage, a retrotranscribed Hnrnph1 was inserted next to the BGW motif and upstream of Gla, creating the Hnrnph2 retrogene; (iii) The region containing the co-opted BGW motif and Hnrnph2 suffered a duplication and HAL1b and L1ME-like retrotransposons were inserted in the vicinity; (iv) The BGW motif and the ORF created by the insertion of the retrotransposons conformed the proto-Bex/Tceal with the YY1 binding site from HAL1b preserved; (v) The BGW motif and the YY1 binding site of a Bex/Tceal gene duplicated upstream of a retrocopy of the Armc10 gene, giving rise to the ArmcX ancestral gene; and (vi) The Bex/Tceal and ArmcX gene families expanded forming the BGW cluster before the diversification of the placental lineage.

Regarding their expression, according to the available data from adult organs, most genes present a tissue-enriched pattern, with the brain be the organ showing the highest expression levels in most of the paralogs. Navas-Pérez et al. (2020) also reported an analysis of the expression via *in situ* hybridization during mouse development, where they observed that especially the Bex genes were highly and widely expressed with Bex3 strongly expressed in the nervous system.

Functional analyses of Bex/Tceal genes have been performed *in vitro*, where Bex3 for example was linked to neuronal physiology, and also through mutant mouse lines. The homozygous mutant lines showed subtle facial differences to the naked eye due to cranial morphology aberrances. The cerebellum and brain showed a reduction in size a fact that may be linked to the behavioral defects observed in the mutant mice. Particularly, the mutants showed impairment in social interactions, nest building, working memory, and object recognition memory. This mutant phenotype coupled with the reported physical interaction of BEX3 with TSC1 (Yasui et al., 2007) could mean that Bex3 is preventing the TSC1/2 complex from interacting with mTORC2, inhibiting this pathway. Thus, Bex3 would be fine-tuning the regulation of the mTOR pathway and its deregulation was suggested to be related to autism spectrum disorder in humans (Ganesan et al., 2019).

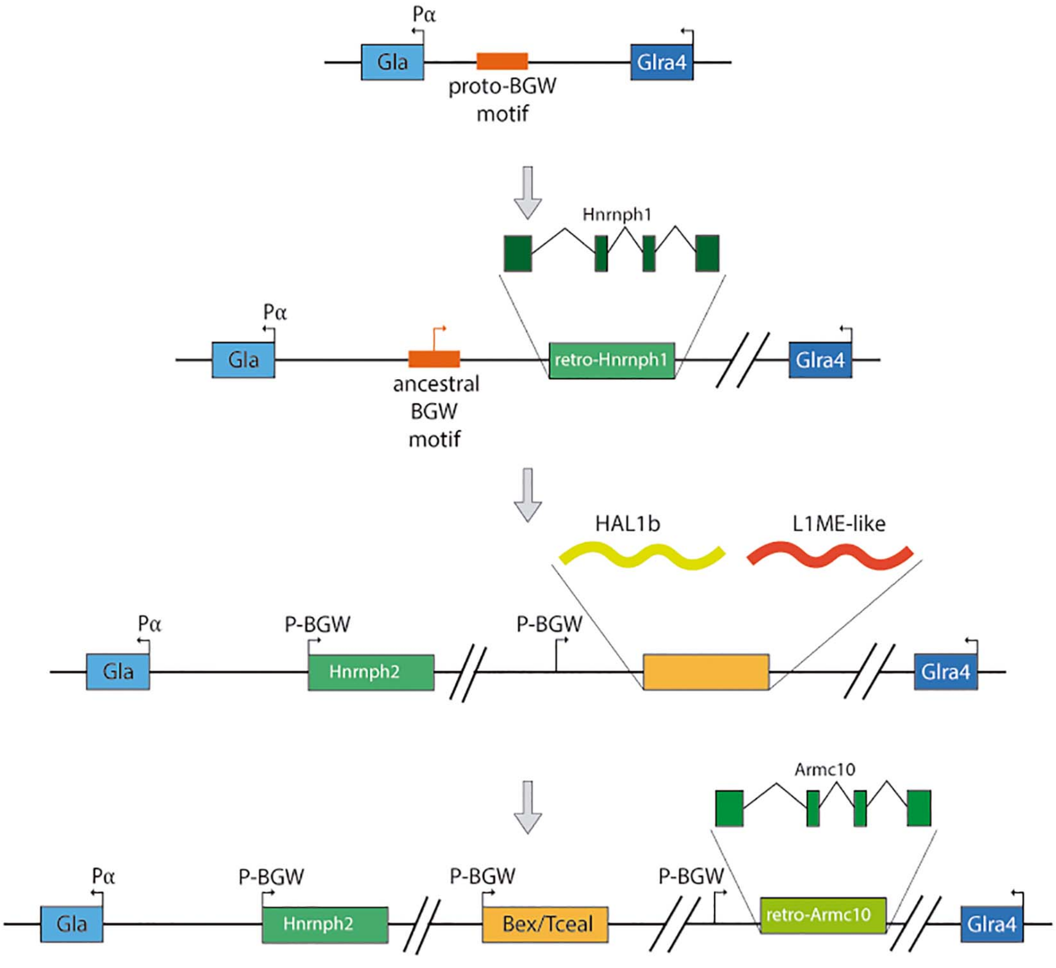


FIGURE 4 | Proposed steps for the Bex/Tceal cluster formation.

Some of the structural features present in BEX and TCEAL families have been detected in the ancestral transposon HAL1b (Navas-Pérez et al., 2020), which indicates that they were preserved along the domestication process., and a positive selection signature can be found in particular cases, suggesting than ancestral genes went through an adaptative process before the diversification of placental mammals. The differences in the neocortex complexity between eutherian and non-eutherian mammals (Cheung et al., 2010) may thus be linked to genomic novelties that emerged during this transition affecting neural development, and maybe, to the appearance of the Bex/Tceal cluster.

The ability to affect neural proliferation, however, seems to have been acquired at some point during the formation of the cluster and not be intrinsic to the eutherian Bex/Tceal ancestor. This was proved electroporating a synthetically reconstructed version of the ancestral Bex/Tceal protogene, as well as the murine Bex3 and Tceal7 into the neural tube of chicken embryos to serve as a non-eutherian vertebrate environment. Expression of Bex3 and Tceal7 generated a noticeable increase in cell proliferation in the embryonic neural tube, similar to what had been reported in mammalian cell cultures (Calvo et al., 2015). Meanwhile, the reconstructed protogene was not able to trigger cellular proliferation, which could mean that this ability was gained afterward its formation.

FINAL REMARKS: NEW GENES, OLD GENOMES, INNOVATIONS, ADAPTATIONS AND, MAYBE, VIRAL FIGHTING SOLUTIONS

Here, we have reviewed a series of cases in which new born genes have been instrumental for the emergence of novelties, some of them shaping a whole taxon. The *de novo* genes, irrespectively if they have been gained by horizontal transfer event or have been formed within the genome, are an engine of evolution, providing new tools for the regulatory networks of extant genomes. Therefore, disruptive novelties may emerge when new genes integrate into old genomes. In the particular case of vertebrates, it is clear that new exons and whole genes were born by recruitment of non-coding DNA and domestication of transposons and virus, that continuously invaded the genome trough vertebrate evolution. Particularly noticeable are some periods linked to the appearance of remarkable innovations, such as the origin of the adaptative immune system, the origin of the placenta, or the deployment of a well-developed neural cortex. Further, in the time of COVID pandemics by SARS-CoV-2, it is tempting to speculate, and timely to investigate, if one of the most successful lineages of mammals, the Chiroptera, is particularly remarkable with regards to *de novo* genes. Bats are particularly more tolerant to viral infections than most mammals, including humans. As somewhat expected for their high resistance to viral infections, the bat genome shows gene expansions and deletions related to the immune system gene network (Jebb et al., 2020) but intriguingly also increasing numbers and high diversity of endogenous viral elements[1] and extremely variable numbers and types of transposon remnants, often bat family or bat species-specific (Jebb et al., 2020). All this together is suggestive of a high level of recent events of virus and transposon invasions in the bat lineage. Whether those invasions helped, in fact, the deployment of virus resistance, in an unexplored but exciting similarity to the original function of the CRISPR/Cas9 system in bacteria, the primitive procaryotic acquired adaptative immunity system, is something that remains to be investigated, with the ambition to learn how to fight, in humans, virus related diseases.

AUTHOR CONTRIBUTIONS

CH-Ú and JG-F: conceptualization, writing, review, and editing. JG-F funding acquisition. Both authors contributed to the article and approved the submitted version.

FUNDING

This article elaboration was funded by grant BFU2017-86152-P (Ministerio de Ciencia, Innovación y Universidades, Spanish Government) to JG-F.

ACKNOWLEDGMENTS

We wish to thanks Michal Zawiska and Claudia Pérez-Calles for critical reading and advice.

---

[1] https://genome-public.pks.mpg.de

REFERENCES

Agassiz, L. (1874). *Evolution and Permanence of Type*. United Kingdom: Universidad de Oxford.

Ashley, J., Cordy, B., Lucia, D., Fradkin, L. G., Budnik, V., and Thomson, T. (2018). Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. *Cell* 172, 262–274.e11. doi: 10.1016/j.cell.2017.12.022

Baguñà, J., and Garcia-Fernàndez, J. (2003). Evo-Devo: the Long and Winding Road. *Int. J. Dev. Biol.* 47, 705–713. doi: 10.1387/ijdb.14756346

Bateson, W. (1922). Evolutionary faith and modern doubts. *Nature* 109, 553–556. doi: 10.1038/109553a0

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., and Noll, M. (1986). Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell* 47, 1033–1040. doi: 10.1016/0092-8674(86)90818-4

Calvo, L., Anta, B., López-Benito, S., Martín-Rodriguez, C., Lee, F. S., Pérez, P., et al. (2015). Bex3 dimerization regulates NGF-dependent neuronal survival and differentiation by enhancing trkA gene transcription. *J. Neurosci.* 35, 7190–7202. doi: 10.1523/JNEUROSCI.4646-14.2015

Carmona, L. M., and Schatz, D. G. (2017). New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. *FEBS J.* 284, 1590–1605. doi: 10.1111/febs.13990

Cavalli, G., and Misteli, T. (2013). Functional implications of genome topology. *Nat. Struct. Mol. Biol.* 20, 290–299. doi: 10.1038/nsmb.2474

Cheung, A. F. P., Kondo, S., Abdel-Mannan, O., Chodroff, R. A., Sirey, T. M., Bluy, L. E., et al. (2010). The subventricular zone is the developmental milestone of a 6-layered neocortex: comparisons in metatherian and eutherian mammals. *Cereb. Cortex* 20, 1071–1081. doi: 10.1093/cercor/bhp168

Cornelis, G., Funk, M., Vernochet, C., Leal, F., Tarazona, O. A., Meurice, G., et al. (2017). An endogenous retroviral envelope syncytin and its cognate receptor identified in the viviparous placental Mabuya lizard. *Proc. Natl. Acad. Sci. U. S. A.* 114, E10991–E11000. doi: 10.1073/pnas.1714590114

Cosby, R. L., Judd, J., Zhang, R., Zhong, A., Garry, N., Pritham, E. J., et al. (2021). Recurrent evolution of vertebrate transcription factors by transposase capture. *Science* 371:eabc6405. doi: 10.1126/science.abc6405

Dahl, E., Koseki, H., and Balling, R. (1997). Pax genes and organogenesis. *BioEssays* 19, 755–765. doi: 10.1002/bies.950190905

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex (1st ed.)*. London: John Murray.

Delaye, L., DeLuna, A., Lazcano, A., and Becerra, A. (2008). The origin of a novel gene through overprinting in *Escherichia coli*. *BMC Evol. Biol.* 8:31. doi: 10.1186/1471-2148-8-31

Deline, B., Greenwood, J. M., Clark, J. W., Puttick, M. N., Peterson, K. J., and Donoghue, P. C. J. (2018). Evolution of metazoan morphological disparity. *Proc. Natl. Acad. Sci. U. S. A.* 115, E8909–E8918. doi: 10.1073/pnas.1810575115

Drost, H.-G., Janitza, P., Grosse, I., and Quint, M. (2017). Cross-kingdom comparison of the developmental hourglass. *Curr. Opin. Genet. Dev.* 45, 69–75. doi: 10.1016/j.gde.2017.03.003

Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development* 1994, 135–142. doi: 10.1242/dev.1994.Supplement.135

Eldredge, N., and Gould, S. (1972). "Punctuated Equilibria: an Alternative to Phyletic Gradualism," In *Models in Paleobiology*, ed. T.J.M. Schopf (San Francisco: Freeman, Cooper and Co), 82–115.

Fromm, B., Billipp, T., Peck, L. E., Johansen, M., Tarver, J. E., King, B. L., et al. (2015). A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu. Rev. Genet.* 49, 213–242. doi: 10.1146/annurev-genet-120213-092023

Ganesan, H., Balasubramanian, V., Iyer, M., Venugopal, A., Subramaniam, M. D., Cho, S. G., et al. (2019). mTOR signalling pathway - A root cause for idiopathic autism?. *BMB Rep.* 52, 424–433. doi: 10.5483/BMBRep.2019.52.7.137

Garcia-Fernàndez, J. (2005). The genesis and evolution of homeobox gene clusters. *Nat. Rev. Genet.* 6, 881–892. doi: 10.1038/nrg1723

Garcia-Fernàndez, J., Marfany, G., Baguñà, J., and Saló, E. (1993). Infiltration of mariner elements. *Nature* 364:109. doi: 10.1038/364109a0

Gracheva, E. O., Cordero-Morales, J. F., González-Carcacía, J. A., Ingolia, N. T., Manno, C., Aranguren, C. I., et al. (2011). Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. *Nature* 476, 88–92. doi: 10.1038/nature10245

Haeckel, E. (1866). *Generelle Morphologie der Organismen*. Berlin: G. Reimer, 1036. doi: 10.1515/9783110848281

Heimberg, A. M., Sempere, L. F., Moy, V. N., Donoghue, P. C. J., and Peterson, K. J. (2008). MicroRNAs and the advent of vertebrate morphological complexity. *Proc. Natl. Acade. Sci. U. S. A.* 105, 2946–2950. doi: 10.1073/pnas.0712259105

Hoshiyama, D., Iwabe, N., and Miyata, T. (2007). Evolution of the gene families forming the Pax/Six regulatory network: isolation of genes from primitive animals and molecular phylogenetic analyses. *FEBS Lett.* 581, 1639–1643. doi: 10.1016/j.febslet.2007.03.027

Huxley, J. (1942). *Evolution, the Modern Synthesis*. Pennsylvania: Universidad Estatal de Pensilvania.

Irie, N., and Kuratani, S. (2011). Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat. Commun.* 2, 1–8. doi: 10.1038/ncomms1248

Irie, N., and Sehara-Fujisawa, A. (2007). The vertebrate phylotypic stage and an early bilaterian-related stage in mouse embryogenesis defined by genomic information. *BMC Biol.* 5:1. doi: 10.1186/1741-7007-5-1

Irimia, M., Denuc, A., Burguera, D., Somorjai, I., Martin-Duran, J. M., Genikhovich, G., et al. (2011). Stepwise assembly of the Nova-regulated alternative splicing network in the vertebrate brain. *Proc. Natl. Acad. Sci. U. S. A.* 108, 5319–5324. doi: 10.1073/pnas.1012333108

Irimia, M., Denuc, A., Ferran, J. L., Pernaute, B., Puelles, L., Roy, S. W., et al. (2012). Evolutionarily conserved A-to-I editing increases protein stability of the alternative splicing factor Nova1. *RNA Biol.* 9, 12–21. doi: 10.4161/rna.9.1.18387

Jacob, F. (1977). Evolution and tinkering. *Science* 196, 1161–1166. doi: 10.1126/science.860134

Jebb, D., Huang, Z., Pippel, M., Hughes, G. M., Lavrichenko, K., Devanna, P., et al. (2020). Six reference-quality genomes reveal evolution of bat adaptations. *Nature* 583, 578–584. doi: 10.1038/s41586-020-2486-3

Jimenez-Delgado, S., Pascual-Anaya, J., and Garcia-Fernandez, J. (2009). Implications of duplicated cis-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty. *Brief. Funct. Genomic. Proteomic.* 8, 266–275. doi: 10.1093/bfgp/elp029

Katzourakis, A., and Gifford, R. J. (2010). Endogenous viral elements in animal genomes. *PLoS Genet.* 6:1001191. doi: 10.1371/journal.pgen.1001191

Kowalewski, A. O. (1866). Entwicklungsgeschichte der einfachen Ascidien. *Mem I'Acad St Petersbourg Ser.* 7, 1–9.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi: 10.1038/35057062

Lavialle, C., Cornelis, G., Dupressoir, A., Esnault, C., Heidmann, O., Vernochet, C., et al. (2013). Paleovirology of "syncytins", retroviral env genes exapted for a role in placentation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120507. doi: 10.1098/rstb.2012.0507

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* 276, 565–570. doi: 10.1038/276565a0

Liscovitch-Brauer, N., Alon, S., Porath, H. T., Elstein, B., Unger, R., Ziv, T., et al. (2017). Trade-off between Transcriptome Plasticity and Genome Evolution in Cephalopods. *Cell* 169, 191–202.e11. doi: 10.1016/j.cell.2017.03.025

Liu, J., Frochaux, M., Gardeux, V., Deplancke, B., and Robinson-Rechavi, M. (2020). Inter-embryo gene expression variability recapitulates the hourglass pattern of evo-devo. *BMC Biol.* 18:129. doi: 10.1186/s12915-020-00842-z

Mayr, E., and Provine, W. B. (1980). *The Evolutionary Synthesis?: Perspectives On The Unification Of Biology*. Cambridge: Harvard University Press.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* 308, 428–433. doi: 10.1038/308428a0

Navas-Pérez, E., Vicente-García, C., Mirra, S., Burguera, D., Fernàndez-Castillo, N., Ferrán, J. L., et al. (2020). Characterization of an eutherian gene cluster generated after transposon domestication identifies Bex3 as relevant for advanced neurological functions. *Genome Biol.* 21:267. doi: 10.1186/s13059-020-02172-3

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin: Springer. doi: 10.1007/978-3-642-86659-3

Paixão-Côrtes, V. R., Salzano, F. M., and Bortolini, M. C. (2015). Origins and evolvability of the PAX family. *Semin. Cell Dev. Biol.* 44, 64–74. doi: 10.1016/j.semcdb.2015.08.014

Parrish, N. F., and Tomonaga, K. (2018). A Viral (Arc)hive for Metazoan Memory. *Cell* 172, 8–10. doi: 10.1016/j.cell.2017.12.029

Pascual-Carreras, E., Marin-Barba, M., Herrera-Úbeda, C., Font-Martín, D., Eckelt, K., de Sousa, N., et al. (2020). Planarian cell number depends on Blitzschnell, a novel gene family that balances cell proliferation and cell death. *Development* 147:dev.184044. doi: 10.1242/dev.184044

Pastuzyn, E. D., Day, C. E., Kearns, R. B., Kyrke-Smith, M., Taibi, A. V., McCormick, J., et al. (2018). The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell* 172, 275–288.e18. doi: 10.1016/j.cell.2017.12.024

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Raff, R. A. (1996). *The Shape of Life*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226256573.001.0001

Roberts, R. M., Green, J. A., and Schulz, L. C. (2016). The evolution of the placenta. *Reproduction* 152, R179–R189. doi: 10.1530/REP-16-0325

Rose, K. M., Hirsch, V. M., and Bouamr, F. (2020). Budding of a retrovirus: Some assemblies required. *Viruses* 12:1188. doi: 10.3390/v12101188

Saxton, R. A., and Sabatini, D. M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. *Cell* 168, 960–976. doi: 10.1016/j.cell.2017.02.004

Scott, M. P., and Weiner, A. J. (1984). Structural relationships among genes that control development: Sequence homology between the antennapedia, ultrabithorax, and fushi tarazu loci of Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* 81, 4115–4119. doi: 10.1073/pnas.81.13.4115

Sha, M., Lee, X., Li, X., Li, X., Veldman, G. M., Finnerty, H., et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403, 785–789. doi: 10.1038/35001608

Simakov, O., and Kawashima, T. (2017). Independent evolution of genomic characters during major metazoan transitions. *Dev. Biol.* 427, 179–192. doi: 10.1016/j.ydbio.2016.11.012

Singh, U., and Syrkin Wurtele, E. (2020). How new genes are born. *Elife* 9:e55136. doi: 10.7554/eLife.55136

Slack, J. M. W., Holland, P. W. H., and Graham, C. F. (1993). The zootype and the phylotypic stage. *Nature* 361, 490–492. doi: 10.1038/361490a0

Stanley, S. M. (1973). An Ecological Theory for the Sudden Origin of Multicellular Life in the Late Precambrian. *Proc. Natl. Acad. Sci.U. S. A.* 70, 1486–1489. doi: 10.1073/pnas.70.5.1486

Van Oss, S. B., and Carvunis, A.-R. R. (2019). De novo gene birth. *PLoS Genet.* 15:e1008160. doi: 10.1371/journal.pgen.1008160

von Baer, K. E. (1828). *Über Entwickelungsgeschichte der Thiere. Beobachtung und Reflexion*. Königsberg: Bei den Gebrüdern Bornträger.

Wagner, G. P. (2014). *Homology, Genes, and Evolutionary Innovation*. Princeton: Princeton University Press. doi: 10.1515/9781400851461

Wang, K. C., and Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Mol. Cell* 43, 904–914. doi: 10.1016/j.molcel.2011.08.018

Wang, Y. T., Pan, Y. J., Cho, C. C., Lin, B. C., Su, L. H., Huang, Y. C., et al. (2010). A novel pax-like protein involved in transcriptional activation of cyst wall protein genes in Giardia lamblia. *J. Biol. Chem.* 285, 32213–32226. doi: 10.1074/jbc.M110.156620

Watson, J. D., and Crick, F. H. C. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737–738. doi: 10.1038/171737a0

Winter, E. E., and Ponting, C. P. (2005). Mammalian BEX, WEX and GASP genes: coding and non-coding chimaerism sustained by gene conversion events. *BMC Evol. Biol.* 5:54. doi: 10.1186/1471-2148-5-54

Yasui, S., Tsuzaki, K., Ninomiya, H., Floricel, F., Asano, Y., Maki, H., et al. (2007). The TSC1 gene product hamartin interacts with NADE. *Mol. Cell. Neurosci.* 35, 100–108. doi: 10.1016/J.MCN.2007.02.007

Zhang, Y., Cheng, T. C., Huang, G., Lu, Q., Surleac, M. D., Mandell, J. D., et al. (2019). Transposon molecular domestication and the evolution of the RAG recombinase. *Nature* 569, 79–84. doi: 10.1038/s41586-019-1093-7

Conflict of Interest: 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 © 2021 Herrera-Úbeda and Garcia-Fernàndez. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 15 July 2021
doi: 10.3389/fcell.2021.696875

A Preliminary Single-Cell RNA-Seq Analysis of Embryonic Cells That Express *Brachyury* in the Amphioxus, *Branchiostoma japonicum*

Noriyuki Satoh[1]*, Hitoshi Tominaga[1], Masato Kiyomoto[2], Kanako Hisata[1], Jun Inoue[3] and Koki Nishitsuji[1]

[1] Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan, [2] Tateyama Marine Laboratory, Marine and Coastal Research Center, Ochanomizu University, Chiba, Japan, [3] Atmosphere and Ocean Research Institute, University of Tokyo, Chiba, Japan

Among chordate taxa, the cephalochordates diverged earlier than urochordates and vertebrates; thus, they retain unique, primitive developmental features. In particular, the amphioxus notochord has muscle-like properties, a feature not seen in urochordates or vertebrates. Amphioxus contains two *Brachyury* genes, *Bra1* and *Bra2*. *Bra2* is reportedly expressed in the blastopore, notochord, somites, and tail bud, in contrast to a low level of *Bra1* expression only in notochord. To distinguish the expression profiles of the two *Brachyury* genes at the single-cell level, we carried out single-cell RNA-seq (scRNA-seq) analysis using the amphioxus, *Branchiostoma japonicum*. This scRNA-seq analysis classified *B. japonicum* embryonic cells into 15 clusters at developmental stages from midgastrula to early swimming larva. *Brachyury* was expressed in cells of clusters 4, 5, 8, and 9. We first confirmed that cluster 8 comprises cells that form somites since this cluster specifically expresses four myogenic factor genes. Cluster 9 contains a larger number of cells with high levels of *Bra2* expression and a smaller number of cells with *Bra1* expression. Simultaneous expression in cluster 9 of tool-kit genes, including *FoxA*, *Goosecoid*, and *hedgehog*, showed that this cluster comprises cells that form the notochord. Expression of *Bra2*, but not *Bra1*, in cells of clusters 4 and 5 at the gastrula stage together with expression of *Wnt1* and *Caudal* indicates that clusters 4 and 5 comprise cells of the blastopore, which contiguously form the tail bud. In addition, *Hox1, Hox3*, and *Hox4* were highly expressed in *Bra2*-expressing clusters 4, 5, 8, and 9 in a temporally coordinated manner, suggesting roles of anterior Hox genes in specification of mesodermal organs, including somites, notochord, and tail bud. This scRNA-seq analysis therefore highlights differences between the two *Brachyury* genes in relation to embryonic regions in which they are expressed and their levels

OPEN ACCESS

Edited by:
Salvatore D'Aniello,
Zoological Station Anton Dohrn, Italy

Reviewed by:
Manuel Irimia,
Centre for Genomic Regulation (CRG), Spain
Carmen Andrikou,
University of Bergen, Norway

***Correspondence:**
Noriyuki Satoh
norisky@oist.jp

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *22 April 2021*
Accepted: *08 June 2021*
Published: *15 July 2021*

Citation:
Satoh N, Tominaga H, Kiyomoto M, Hisata K, Inoue J and Nishitsuji K (2021) A Preliminary Single-Cell RNA-Seq Analysis of Embryonic Cells That Express Brachyury in the Amphioxus, Branchiostoma japonicum. Front. Cell Dev. Biol. 9:696875. doi: 10.3389/fcell.2021.696875

of expression. *Bra2* is the ancestral *Brachyury* in amphioxus, since expression in the blastopore is shared with other deuterostomes. On the other hand, *Bra1* is a duplicate copy and likely evolved a supplementary function in notochord and somite formation in the *Branchiostoma* lineage.

Keywords: scRNA-seq analysis, amphioxus embryos, two *Brachyury* genes, notochord, somites, blastopore, tail bud, *Bra2* orthology

INTRODUCTION

The origin and evolution of chordates are two of the most intriguing evo-devo research subjects of metazoans (Holland et al., 2015; Lowe et al., 2015; Satoh, 2016; Gee, 2018). Among the three chordate taxa, cephalochordates, urochordates (tunicates), and vertebrates, cephalochordates were the first group to diverge (Bourlat et al., 2006; Delsuc et al., 2006; Putnam et al., 2008); thus, they exhibit unique features during embryogenesis, while sharing others with echinoderms and hemichordates, and still others with urochordates and vertebrates (Whittaker, 1997: Yu et al., 2007; Bertrand and Escriva, 2011; Holland et al., 2015).

In particular, the developmental mode and properties of the cephalochordate notochord are unique, compared with those of tunicates and vertebrates (Satoh et al., 2012). First, the cephalochordate notochord is formed at roughly the neurulation stage by "pouching-off" from the dorsal region of the archenteron, like somites, which are also pinched off from both the left and right sides of the archenteron (Conklin, 1932; Hirakow and Kajita, 1994; **Figure 1**). This contrasts with urochordate and vertebrate notochords, which are formed by convergent extensions of precursor cells that are bilaterally positioned in the early embryo (Munro and Odell, 2002). Second, the cephalochordate notochord has muscle-like properties and possesses myofibrils (Ruppert, 1997; Suzuki and Satoh, 2000; Urano et al., 2003). This also contrasts with urochordate and vertebrate notochords, which do not possess any muscle-like properties, except for the posterior part of the *Ciona* embryonic notochord (Cao et al., 2019). Vacuolation within notochord cells provides both stiffness and an increase in cell volume in ascidians and vertebrates (Jiang and Smith, 2007).

We are interested in a T-box family transcription factor gene, *Brachyury*, and its role in formation of the notochord during chordate evolution (Satoh et al., 2012; Satoh, 2016). In non-chordate deuterostomes, *Brachyury* is expressed in the blastopore and stomodeum of early embryos and functions in morphogenetic movement of archenteron invagination and mouth invagination (Tagawa et al., 1998; Gross and McClay, 2001). In chordates, *Brachyury* is expressed in the notochord, and functions in development of this chordate-specific organ. We consider the former primary and the latter secondary expression and function of *Brachyury* (Satoh et al., 2012). Thus, the acquisition of secondary expression and function of *Brachyury* is a foundational research subject to decipher the origin and evolution of chordates.

As mentioned above, the cephalochordate notochord is unique from an evo-devo perspective. In addition, in contrast to the single copy present in genomes of deuterostomes, *Brachyury* was tandemly duplicated into *Bra1* and *Bra2* in cephalochordate genomes (Holland et al., 1995; Inoue et al., 2017). These genes are expressed in the blastopore of gastrula, notochord, somites, and the tail bud of neurulae (Holland et al., 1995; Terazawa and Satoh, 1997; **Figure 1**). However, due to high levels of sequence similarity, previous whole-mount *in situ* hybridization (WMISH) analyses failed to distinguish the expression profiles of *Bra1* and *Bra2.* Recently, Yuan et al. (2020) tackled this question using WMISH with probes from the divergent untranslated regions specific to each of the genes. They showed that the zygotic expression level of *Bra2* is much higher than that of *Bra1,* and that *Bra2* is highly expressed in the blastopore, tail bud, and notochord, while *Bra1* is weakly transcribed only in the notochord. In addition, a heterogenic transplantation assay of *cis*-regulatory sequences of *Bra1* and *Bra2* into the *Ciona* embryonic system demonstrated that the 5' upstream sequence of *Bra2* contains higher enhancer activity in both the notochord and somites, compared to that of the 5' upstream sequence of *Bra1* in the notochord (Tominaga et al., 2018). These results suggest that both *Bra1* and *Bra2* are *Brachyury* orthologs and that *Bra1* has diverged more. Nevertheless, many questions remain to be answered in order to understand evo-devo features of cephalochordate *Brachyury* and the notochord.

The recently developed technology of single-cell RNA sequencing (scRNA-seq) constitutes a powerful tool to categorize genes expressed in constituent cells of embryos, tissues, or organs on a cell-by-cell basis (Cao et al., 2019; Sharma et al., 2019; Stuart and Satija, 2019; Foster et al., 2020). We employed scRNA-seq analyses of amphioxus embryos to clarify how expression levels of *Bra1* and *Bra2* differ in embryonic regions.

RESULTS

Clustering of Embryonic Cells

Branchiostoma japonicum embryos were cultured and collected for scRNA-seq analysis using the 10 × Genomics platform. We selected six embryonic stages from midgastrula to early swimming larva (**Supplementary Table 1**), on the basis of dynamic features of cell fate changes anticipated by previous studies (Bertrand and Escriva, 2011; Marletaz et al., 2018). Datasets from all six stages were integrated and clustered to identify cells with similar or different gene expression profiles across early development using Seurat (Butler et al., 2018; Stuart et al., 2019). In total, transcriptomes of 14,016 cells were included in this analysis. Upon integration of the datasets and visualization with UMAP (Uniform Manifold Approximation and Projection)

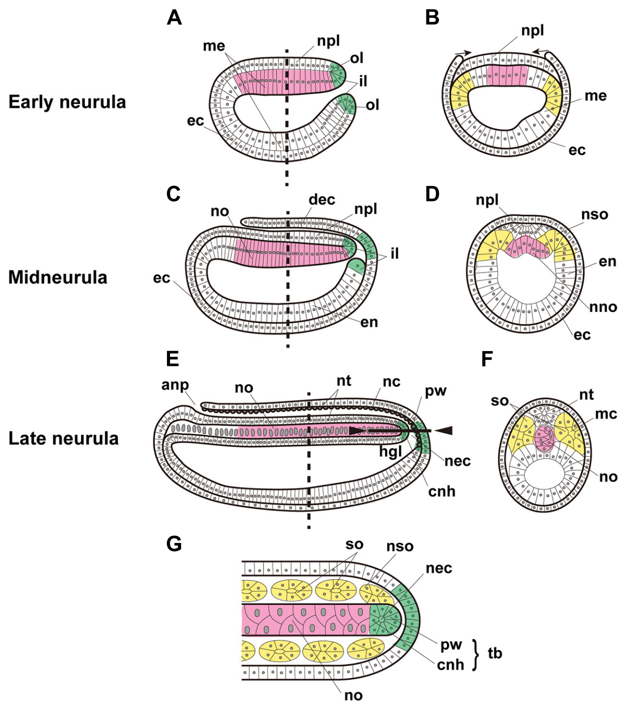


FIGURE 1 | Expected expression of *Brachyury* genes in the notochord (magenta), somites (yellow), and tail bud (green) of amphioxus embryos. Early neurula, **(A)** sagittal section, and **(B)** cross section at the location shown by the dashed line in panel **(A)**. Arrows indicate the migration of ectoderm across the neural plate. Midneurula, **(C)** sagittal section, and **(D)** cross section at the level shown by the dashed line in panel **(C)**. Cross section showing somites and notochord beginning to form. Late neurula, **(E)** sagittal section, **(F)** cross section at the level shown by the dashed line in panel **(E)**, and **(G)** frontal section of the region indicated by the solid line between arrowheads in panel **(E)**. In alphabetical order, abbreviations are as follows: anp, anterior neuropore; cnh, chordoneural hinge; dec, dorsal ectoderm; ec, ectoderm; en, endoderm; hgl, hindgut lumen; il, inner lip of blastopore; mc, myocoel; me, mesendoderm; nc, neurocoel; nec, neurenteric canal; nno, nascent notochord; no, notochord; npl, neural plate; nso, nascent somite; nt, neural tube; ol, outer lip of blastopore; pw, posterior wall of neurenteric canal; so, somite; tb, tail bud. The original drawing of embryos by Schubert et al. (2001) is cited.

dimensionality reduction, we identified 15 major cell clusters, numbered 0 to 14, based on marker gene expression (**Figure 2A** and **Supplementary Table 2**). Identification of 15 clusters resulted from the resolution parameter, 0.5, a conservative setting, used in the clustering step of the analysis.

The present characterization of 15 embryonic-cell clusters encountered two problems. The first was the presence of fewer clusters in the midgastrula stage than in the late gastrula and later stages (**Figure 2B**). This might occur because midgastrula consists of only three germ layers, outer ectoderm, inner endoderm, and intermediate mesoderm, and specification/differentiation of embryonic cells may not have occurred for sufficient clustering of cells (**Figure 2B**). Since the gene encoding type 1 keratin, a marker gene for epidermis specification (Petillon et al., 2017), is specifically expressed in cluster 0, this cluster corresponds to ectoderm (red brown area in **Figure 2B**). As described below, clusters 1, 2, 3, 11, and 12 express *FoxA* and *Goosecoid*, markers for endoderm, so these clusters (**Figure 2B**) constitute endoderm. *Brachyury* is expressed in clusters 4, 5, 8, and 9 (see below), indicating that they are mesoderm territory (**Figure 2B**). The number of clusters therefore increased as development proceeded (compare **Figure 2B** to **Figures 2C,D**). The second was the scarceness of cells that constitute each cluster at early swimming larva (**Figure 2G**). It may be that the available scRNA-seq reads were inadequate to cover a large number of constituent cells of larvae at the stage. However, data from the remaining four stages (late gastrula, early neurula, midneurula, and late neurula) contained enough RNA for further characterization of cell clusters that comprise early amphioxus embryos.

A Cluster of Cells With Myogenic Factor Gene Expression: Cluster 8

Spatial expression profiles of developmentally relevant genes, especially for transcription factors and signaling molecules, have been examined extensively in amphioxus embryos by cDNA cloning and WMISH analysis (Holland et al., 1995; Schubert et al., 2001). Expression profiles of most tool-kit genes did not appear specific to certain embryonic regions, but spanned multiple regions, probably due to a low level of specification in amphioxus early embryonic cells. One example of specific spatial expression was seen among myogenic factor (MF) genes, which are expressed exclusively in the developing paraxial mesoderm or somites (Schubert et al., 2003; Aase-Remedios et al., 2020). Since somites are embryonic organs with expected *Brachyury* expression, and since MF genes are good candidates to examine the appropriateness of our clustering method of amphioxus embryonic cells, we first explored cells of clusters that expressed MF genes.

The amphioxus MF gene was duplicated in a specific manner in this lineage, independently of the vertebrate MF gene duplication into four copies, *MyoD, myogenin*, *Myf5*, and *MRF4* (Araki et al., 1996; Schubert et al., 2003; Tan et al., 2014). Recently, Aase-Remedios et al. (2020) reported five MF genes in *Branchiostoma* genomes. Here, we independently examined the number of MF genes in amphioxus. Surveying a newly assembled genome of *B. floridae* (Simakov et al., 2020), we found a cluster of five MF-related genes in scaffold NC_049997.1 of chromosome 19 (**Figure 3Aa**). They are *MDF* (gene model ID, LOC118406741), *MyoD1* (AY154744; LOC118407021), an uncharacterized copy (LOC118407176, tentatively called *MDF-candidate*), *MyoD2* (AY154745; LOC118406750), and a gene for myoblast determination protein (LOC118406791, tentatively called *MDP*) (**Figure 3Aa**). A survey of the *Branchiostoma belcheri* genome (Huang et al., 2014) showed an orthologous cluster with five genes in scaffold NW_017804132.1 (**Figure 3Ab**), *MDF* (AY066009; LOC109480333), *MyoD1* (AB092415; LOC109480329), *MDF-candidate* (AY313170; LOC109480322), *MyoD2* (AB092416; LOC109480315), and *MDP* (LOC109480330) (**Figure 3Ab**). Molecular phylogeny using ORTHOSCOPE (Inoue and Satoh, 2018) showed that (a) *MDF* and *MyoD2* form a clade that is sister to *MyoD1*, and

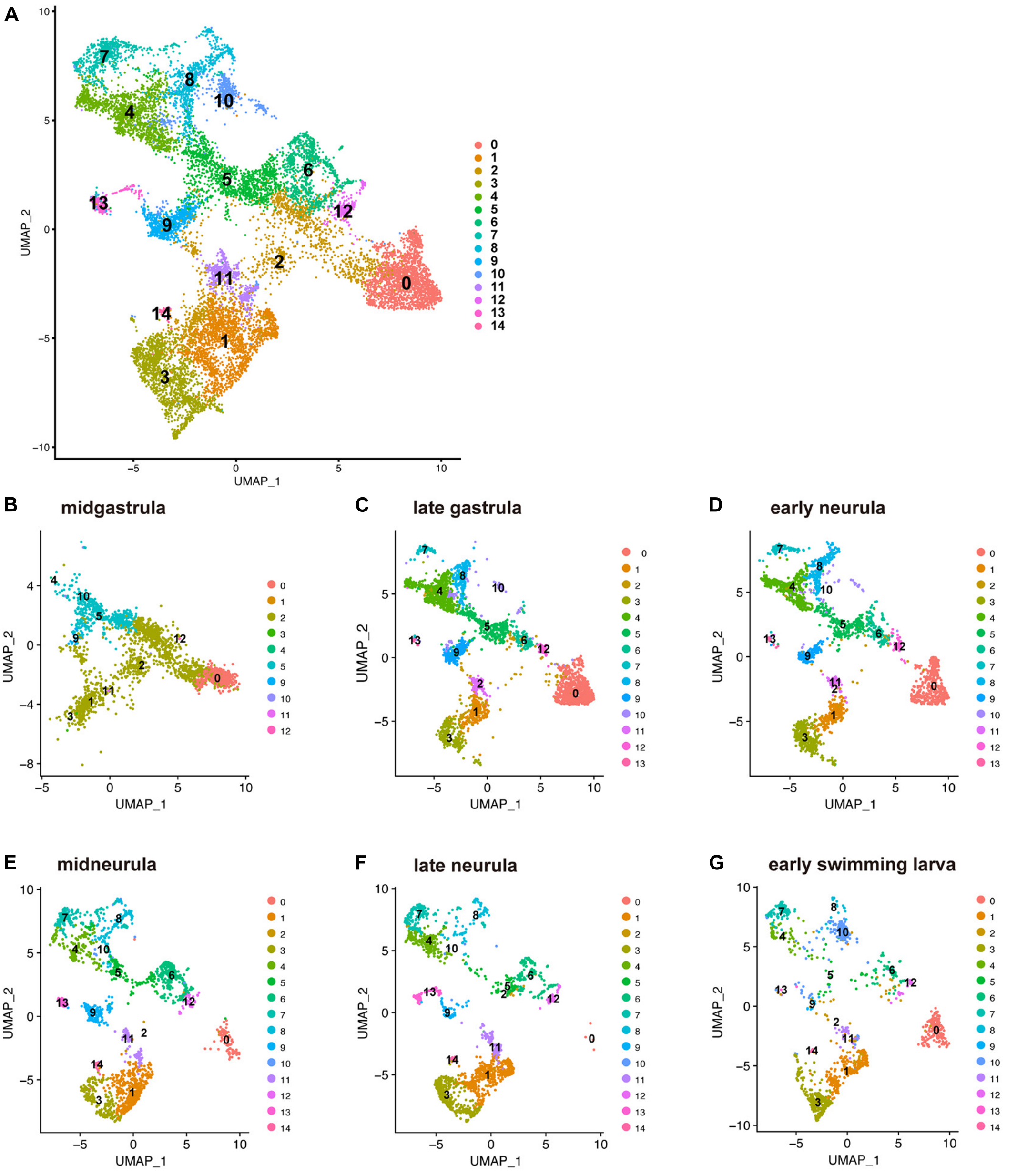


FIGURE 2 | UMAP plots of early developmental stages of *Branchiostoma japonicum*. **(A)** UMAP plots based on a combined dataset of all six stages. **(B)** Midgastrula, **(C)** late gastrula, **(D)** early neurula, **(E)** midneurula, **(F)** late neurula, and **(G)** early swimming larva after alignment. A total of 14,016 single cells are colored by cluster identity. Fifteen populations (clusters 0–14) are detected across the six time points. As discussed in the text, cluster 0 corresponds to ectoderm (red brown area) since a marker gene that encodes type 1 keratin is expressed in the cluster. Clusters 1, 2, 3, 11, and 12 correspond to endoderm (light green and brown area) since *FoxA* and *Goosecoid*, markers for endoderm, are expressed in these clusters. Clusters 4, 5, 8, and 9 correspond to mesoderm (green and blue area), since *Brachyury* is expressed in these clusters. Clusters 6, 7, and 10 may be mesoderm, while clusters 13 and 14 remain to be characterized.

(b) *MDF-candidate* is an early branch of amphioxus MF genes and *MDP* (**Supplementary Figure 1A**). These results coincided well with those of Aase-Remedios et al. (2020); *MDF, MyoD1, MDF-candidate, MyoD2,* and *MDP* correspond to *MRF2b, MRF1, MRF3, MRF2a*, and *MRF4* of Aase-Remedios et al. (2020), respectively (**Figure 3Ab**).

With the five amphioxus MF genes as probes, we carried out scRNA-seq analysis and results obtained are shown in UMPA plots (**Supplementary Figure 2**) and dot plots (**Figure 4**). We found that (a) of the five MF genes, a cell cluster expressing *MDF* (**Supplementary Figure 2A**), *MyoD1* (**Supplementary Figure 2B**), *MyoD2* (**Supplementary Figure 2C**), and *MDP* (**Supplementary Figure 2D**) was evident, but no clusters expressed *MDF-candidate* at a detectable level (data not shown); (b) all four MF genes were expressed in cells that comprise cluster 8 (**Figure 4A** and **Supplementary Figure 2**); (c) the

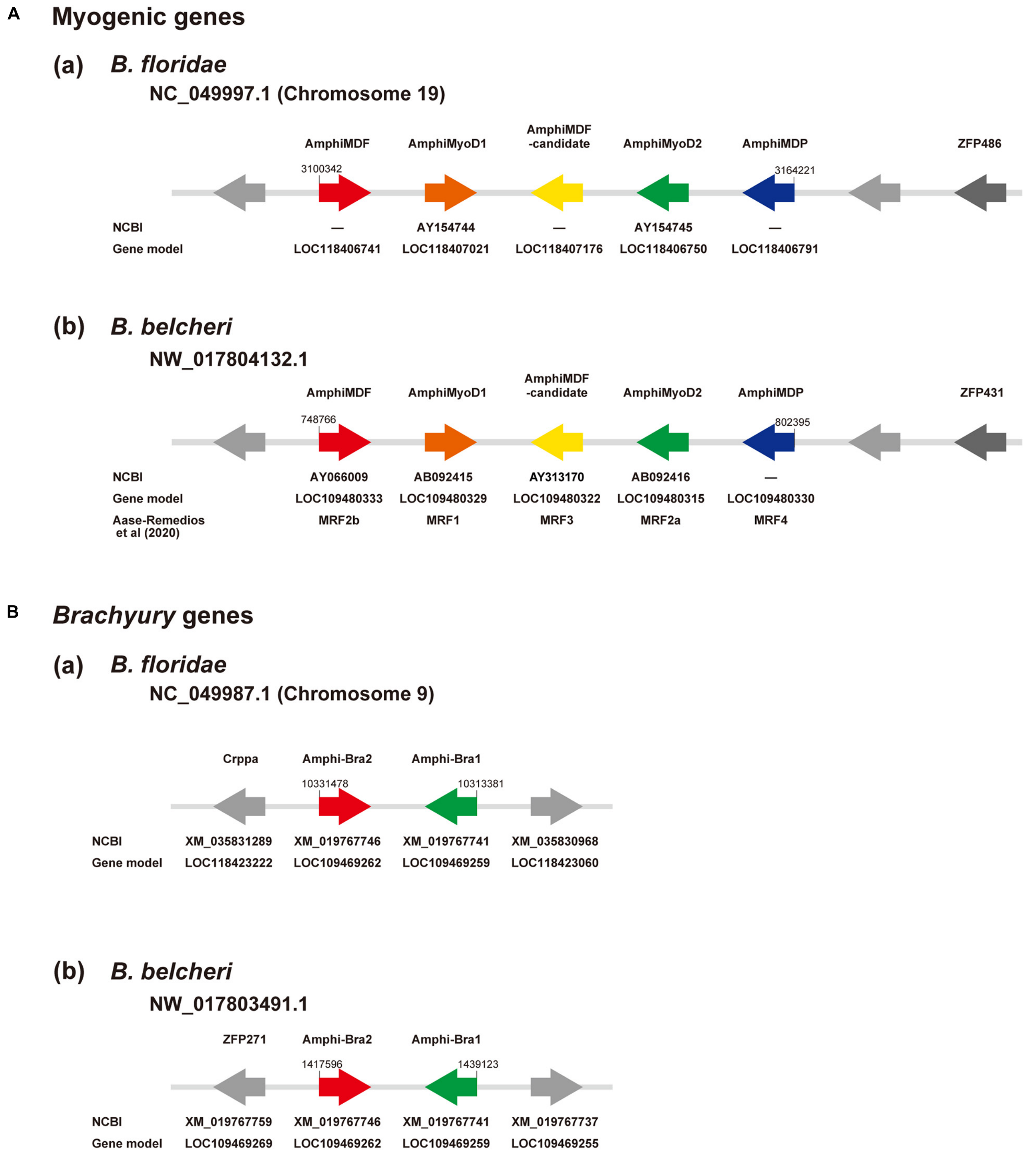

A
Myogenic genes
(a) B. floridae
NC_049997.1 (Chromosome 19)
AmphiMDF
AmphiMyoD1
AmphiMDF -candidate
AmphiMyoD2
AmphiMDP
ZFP486
3100342
3164221
NCBI
—
AY154744
—
AY154745
—
Gene model
LOC118406741
LOC118407021
LOC118407176
LOC118406750
LOC118406791
(b) B. belcheri
NW_017804132.1
AmphiMDF
AmphiMyoD1
AmphiMDF -candidate
AmphiMyoD2
AmphiMDP
ZFP431
748766
802395
NCBI
AY066009
AB092415
AY313170
AB092416
—
Gene model
LOC109480333
LOC109480329
LOC109480322
LOC109480315
LOC109480330
Aase-Remedios et al (2020)
MRF2b
MRF1
MRF3
MRF2a
MRF4
B
Brachyury genes
(a) B. floridae
NC_049987.1 (Chromosome 9)
Crppa
Amphi-Bra2
Amphi-Bra1
10331478
10313381
NCBI
XM_035831289
XM_019767746
XM_019767741
XM_035830968
Gene model
LOC118423222
LOC109469262
LOC109469259
LOC118423060
(b) B. belcheri
NW_017803491.1
ZFP271
Amphi-Bra2
Amphi-Bra1
1417596
1439123
NCBI
XM_019767759
XM_019767746
XM_019767741
XM_019767737
Gene model
LOC109469269
LOC109469262
LOC109469259
LOC109469255

FIGURE 3 | (A) Myogenic factor genes and **(B)** *Brachyury* genes in amphioxus genomes. **(A)** Genomic region containing a cluster of myogenic factor genes in (a) *Branchiostoma floridae* and (b) *B. belcheri*. In scaffold NC_049997.1 of chromosome 19 in *B. floridae* and NW_017804132.1 in *B. belcheri*, there is a cluster that contains five genes for myogenic factors, from left to right, MDF (red), MyoD1 (brown), a candidate myogenic factor gene (yellow), MyoD2 (green), and myoblast determination protein (blue). Two undefined genes are shown in gray and a zinc finger protein gene is shown in dark gray. Arrows indicate the direction of transcription. **(B)** Genomic region containing two *Brachyury* genes, *Bra2* (red) and *Bra1* (green), in (a) *Branchiostoma floridae* scaffold NC_049987.1 of chromosome 9 and (b) *B. belcheri* scaffold NW_017803491. Arrows indicate the direction of transcription.

expression level of *MDF* and *MDP* was higher than those of *MyoD1* and *MyoD2* (**Figures 4A,B**); (d) expression of *MDF* and *MDP* became evident at the late gastrula stage and continued until the late neurula stage (**Figures 4A,B**); (e) expression of *MyoD1* was detected at the early, mid, and late neurula stage (**Figures 4A,B**); and (f) expression of *MyoD2* was detected at

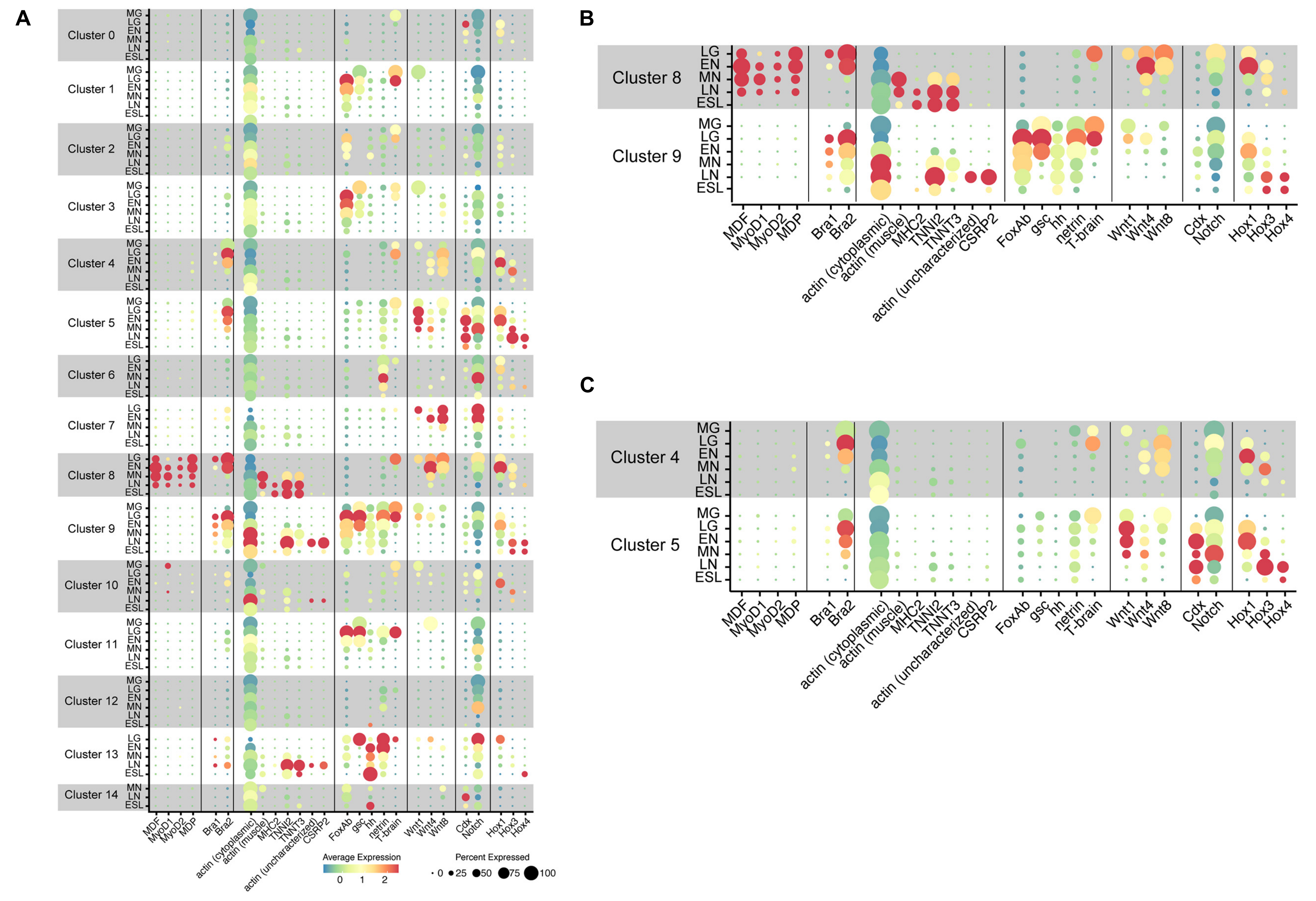


FIGURE 4 | Transcriptome trajectories for 15 clusters comprising early stages of *Branchiostoma japonicum* embryos. Dot size represents the percentage of cells that express genes for transcription factors, signaling molecules, and structural proteins. Dot color shows the averaged level of expression. MG, midgastrula; LG, late gastrula; EN, early neurula; MN, midneurula; LN, late neurula; ESL, early swimming larva. **(A)** A summary figure. **(B)** Enlargement of clusters 8 and 9. **(C)** Enlargement of clusters 4 and 5.

the four earlier stages (**Figures 4A,B**). These results indicate that cells of cluster 8 comprise the developing somites, consistent with results of previous studies showing that amphioxus MF genes are expressed exclusively in embryonic cells that give rise to paraxial muscle or somites at gastrula and neurula stages (Schubert et al., 2003; Urano et al., 2003; Aase-Remedios et al., 2020). Based on their higher expression levels, *MDF* and *MDP* likely function in myogenesis of amphioxus embryos rather than *MyoD1* and *MyoD2*, although this should be further examined experimentally since Aase-Remedios et al. (2020) reported expression of all MF genes in *Branchiostoma* embryos. Nevertheless, the appropriateness of our clustering method was also supported by these results.

Cluster of Cells With *Bra1* and *Bra2* Expression: Cluster 9

As in the case of MF genes, we first examined localization of *Bra1* and *Bra2* in amphioxus genomes and their molecular phylogenetic relationships among deuterostome T-box family members for basic information following scRNA-seq analysis. *Bra1* and *Bra2* were present in scaffold NC_049987.1 of chromosome 9 in the *B. floridae* genome (**Figure 3Ba**) and in scaffold NW_017803491.1 in the *B. belcheri* genome (**Figure 3Bb**). They exhibit a head-to-head orientation, and there were no other T-box family members in this genomic region. Molecular phylogeny using ORTHOSCOPE demonstrated the presence of seven T-box subfamily members in lancelet genomes, TBX20, TBX15/18/22/, TBX1/10, TBX4/5, TBX2/3, TBR1/TBX21, and Bra (**Supplementary Figure 1B**). Duplication into *AmphiBra1* and *AmphiBra2* appears to be an amphioxus-specific event (Inoue et al., 2017). A sister clade to *Bra* clade is *Tbrain*/*Tbx21*. *Tbrain1* is apparently expressed in the dorsal region of invaginating archenteron in gastrulae, and then in the preoral-pit region of larvae (Satoh et al., 2002), suggesting overlapping expression that of *Bra* and *Tbrain1* in gastrulae.

The present scRNA-seq analysis showed that *Bra* is expressed in cells of clusters 4, 5, 8, and 9 (**Figure 4** and **Supplementary Figure 3**). A small number of cells expressed *Bra* weakly in cluster 13 (**Figure 4**), and the property of this cluster should be characterized in future studies. Expression of *Bra2* (**Supplementary Figure 3B**) was much higher than that of *Bra1* (**Supplementary Figure 3A**) both at the level of transcription and in embryonic regions in which the gene is expressed (**Figure 4**). The occurrence of cells with *Bra1* expression was detected in clusters 8 and 9 at late gastrula and in cluster 9 at early neurula and midneurula and became undetectable by the late neurula stage (**Figure 4B**). In addition, cluster 9 contained a larger number of cells with high *Bra2* expression and a smaller number of cells with *Bra1* expression (**Figure 4B**).

As shown in the previous section, cluster 8 pertains to the developing somites. This indicates that cluster 9 comprises cells of the developing notochord. Therefore, at the late gastrula stage, *Bra1* is expressed in both the presumptive notochord and the presumptive paraxial muscle. These results basically coincide with those of Yuan et al. (2020), although they reported no detectable level of *Bra1* expression at the gastrula stage and in somites (cluster 8) (Yuan et al., 2020).

Clusters 8 and 9 Express Muscle Structural Genes

As already mentioned, the cephalochordate notochord has muscle-like properties. Expression of muscle structural genes, including actin and myosin heavy chain, has been reported not only in somites, but also in the notochord (Suzuki and Satoh, 2000; Urano et al., 2003; Inoue and Satoh, 2018). We first examined whether cells of cluster 8 (somites) express muscle structural genes to confirm that this cluster comprises cells for the developing somites. Three genes examined for this purpose were cytoplasmic actin (LOC109483419) (**Supplementary Figure 4A**), paraxial muscle actin (LOC109481400) (**Supplementary Figure 4B**), and myosin heavy chain (MHC2) (LOC109470521) (**Supplementary Figure 4C**). As expected, considerable cytoplasmic actin gene expression was detected in most cell clusters through the six stages (**Figure 4**). In contrast, the gene for paraxial muscle actin was detected in cluster 8 at midneurula and later stages (**Figures 4A,B**), and *MHC2* gene was detected in cluster 8 at late neurula and early larva stages (**Figures 4A,B**). Therefore, the affiliation of cluster 8 with somites was confirmed by expression of paraxial muscle structural genes. Earlier expression of *MDF* and *MDP* than paraxial muscle structural genes (**Figure 4B**) suggests regulatory control of MF genes over transcription of structural genes.

The cephalochordate notochord has been reported to express muscle structural genes, some shared with somites while some others are specific to the notochord (Suzuki and Satoh, 2000; Urano et al., 2003; Inoue and Satoh, 2018). Although detailed identification and characterization of muscle structural genes in the developing cephalochordate notochord using scRNA-seq data are the subject of future research, here we examined four genes, gene for troponin I2 (TNNI2; LOC109469528) (**Supplementary Figure 4D**), for troponin T3 (TNNT3; LOC109481859) (**Supplementary Figure 4E**), for possible notochord-specific actin (not characterized yet) (LOC109482101) (**Supplementary Figure 4F**), and for cysteine- and glycine-rich protein 2 (CSRP2) (LOC109481702) (**Supplementary Figure 4G**). *CSRP2* encodes a group of LIM domain proteins, a cephalochordate homolog of which was reported to express in the developing notochord (Urano et al., 2003). High levels of *troponin I2* and *troponin T3* expression were detected in cluster 8 at late neurula and early swimming larva stages (**Figure 4B**). In addition, high *troponin I2* expression was evident in cluster 9 at late neurula stages (**Figure 4B**). On the other hand, expression of the uncharacterized actin gene and *CSRP2* was detected only in cells of cluster 9 at late neurula (**Figure 4B**). Therefore, this analysis suggests the unique properties of the cephalochordate notochord, in which genes for several muscle structural proteins are likely controlled by *Brachyury*, but not by *MDF*/*MyoD,* since *MDF*/*MyoD* is not expressed in cells of cluster 9.

Characterization of Notochord Cluster 9 With Other Transcription Factor Genes

Previous cDNA cloning studies of amphioxus genes followed by WMISH analysis demonstrated that various genes for transcription factors and signaling molecules are expressed in the embryonic region that gives rise to the notochord. We here examined scRNA-seq data of four genes, namely, *FoxAb* (LOC10948147) (**Supplementary Figure 5A**), *sonic hedgehog* (LOC10947428) (**Supplementary Figure 5B**), *goosecoid* (LOC109470978) (**Supplementary Figure 5C**), and *netrin* (LOC109462789) (**Supplementary Figure 5D**).

FoxA and *Brachyury* function in a coordinated manner to form endomesoderm in deuterostome embryos (e.g., Imai et al., 2006; Oliveri et al., 2006). *Branchiostoma floridae* contains two *FoxA* genes, *FoxAa* and *FoxAb* (Yu et al., 2008). Shimeld (1997) reported that transcripts of *FoxAb* are detected in the blastopore, endoderm, and notochord in amphioxus neurulae. The present scRNA-seq analysis showed that cells with *FoxAb* expression became detectable as early as the midgastrula stage in cluster 9 (**Figure 4** and **Supplementary Figure 5A**). In late gastrula, the high level of expression was found in clusters 1, 2, 3, 9, and 11 (**Figure 4**), and this expression remained in early, mid-, and late neurula (**Figure 4**). This indicates that cells of cluster 9 express *Bra2* and *FoxAb* simultaneously. Clusters 1, 2, 3, and 11 occupy a closely associated region of UMAP plots (**Figure 2**), and because in addition to *FoxAb*, an anterior endodermal marker of amphioxus embryos, *goosecoid* (Onai et al., 2015), is expressed in these clusters (**Figure 4**, see below), cells of clusters 1, 2, 3, and 11 were thought to comprise endoderm and its derivatives.

Goosecoid (*gsc*) encodes a homeodomain-containing transcription factor that is expressed in the vertebrate head organizer and that can induce a secondary axis when expressed ectopically. Neidert et al. (2008) and Onai et al. (2015) reported that the expression of *Amphioxus gsc* is initially localized during gastrulation to the endomesoderm layer of the dorsal lip of the blastopore. Then, *gsc* is expressed in the anterior and dorsal endoderm and in the presumptive notochord that underlies the presumptive central nervous system (Onai et al., 2015). This scRNA-seq analysis showed that *gsc* is expressed in cells of cluster 9 at midgastrula, late gastrula, and early neurula (**Figure 4B** and **Supplementary Figure 5C**), further supporting the notion that cluster 9 consists of cells for the developing notochord. *gsc* expression was also detected in clusters 1, 3, and 11, suggesting that these clusters are derived from endoderm (**Figure 4B** and **Supplementary Figure 5C**).

A single amphioxus hedgehog gene, *hh*, is expressed in the notochord and ventral neural tube, embryonic tissues that express Sonic-type genes in vertebrates (Shimeld, 1999; Ono et al., 2018; Zhu et al., 2020). Our scRNA-seq analysis detected cells with *hh* expression in cluster 9 at late gastrula and later stages (**Figure 4B** and **Supplementary Figure 5B**), indicating that cluster 9 comprises cells that form the notochord. A high level of *hh* expression was detected in cluster 13 at early neurula and later stages (**Figure 4** and **Supplementary Figure 5B**). This suggests that cluster 13 contains cells of neural tube.

An amphioxus *netrin* gene is expressed in midline structures of embryos, including the notochord and floor plate (Shimeld, 2000). This scRNA-seq analysis confirmed that *netrin* is expressed in cells of notochord cluster 9 (**Figure 4B** and **Supplementary Figure 5D**). In addition, *netrin* expression was detected in cells of clusters 6 and 13 at late gastrula and later stages (**Figure 4** and **Supplementary Figure 5D**). As mentioned above, cluster 13 likely involves cells of the nervous system, and *netrin* expression in cluster 13 supports this notion. Cluster 6 likely pertains to mesoderm, but characterization of this cluster will be investigated in future studies.

Tbrain1 is sister to *Bra* in the amphioxus T-box family (**Supplementary Figure 1B**), and this gene is reportedly expressed in the dorsal region of invaginating archenteron in gastrulae (Satoh et al., 2002). This scRNA-seq analysis showed that *Tbrain1* expression occurs in cells of cluster 8 at late gastrula and in cells of cluster 9 at mid- and late gastrula (**Figure 4** and **Supplementary Figure 5E**). This indicates that cells of somite cluster 8 and notochord cluster 9 are derived from the dorsal region of the invaginating archenteron. This overlapping expression of *Tbrain1* with *Bra2* suggests some role of *Tbrain1* in the developing somite and notochord. A high level of *Tbrain1* expression was also detected in cells of clusters 1 and 11 at gastrula (**Figure 4** and **Supplementary Figure 5E**), providing additional support for an association of clusters 1 and 11 with endoderm.

In summary, this scRNA-seq survey of five transcription factor genes almost coincides with results of previous WMISH studies and suggests that cluster 9, with *Brachyur*y expression, comprises cells of the developing notochord. Therefore, these transcription factor genes are likely involved either directly or indirectly in the formation of the notochord in amphioxus embryos.

Co-expression of *Bra2* and *MDF*

The four MF genes, including *MDF*, are specifically expressed in cells of cluster 8 (somite) (**Figure 4**), and *Bra2* is also expressed in cells of clusters 8 (**Figure 4**), showing that presumptive somite cells simultaneously express MF genes and *Brachyury*. To examine co-expression of these genes at the single-cell level, we double-plotted cells with *MDF* expression and those with *Bra2* expression (**Supplementary Figure 6**). At late gastrula, co-expression was detected in most cells of cluster 8 (**Supplementary Figure 6**; LG). On the other hand, at early neurula, cells with only *MDF* expression appeared in the distal portion of the cluster (**Supplementary Figure 6**; EN). As embryos further developed to mid- and late neurula stage, the expression level of *Bra2* decreased (**Figure 4B**) and cluster 8 contained more cells with only *MDF* expression (**Supplementary Figure 6**; MN and LN). These results suggest that at the single-cell level, cells of cluster 8 utilize MF genes and *Brachyury* for specification as presumptive somite cells at the gastrula stage, while later at the neurula stage they use only MF genes for specification.

Characterization of Clusters With *Bra2* Expression: Clusters 4 and 5

As mentioned above, embryonic regions with *Bra2* expression were comparatively broad because this gene is highly expressed not only in notochord cluster 9 and somite cluster 8, but also in clusters 4 and 5 (**Figure 4** and **Supplementary Figure 3**).

Cells of both clusters 4 and 5 began to express *Bra2* at the midgastrula stage and high level of expression was evident at late gastrula, although expression decreased as development proceeded, becoming undetectable at the late neurula stage (**Figure 4C**). Since *Brachyury* is reportedly first expressed in the blastopore region of gastrulae, and then in the tail bud, it is highly likely that clusters 4 and 5 correspond to these regions. Two clusters with *Bra2* expression suggest that the tail bud is composed of two regions with different properties. According to a review of amphioxus tool-kit genes involved in tail bud formation from the blastopore region (Holland, 2002), the tail bud is formed by combinatorial expression of *Brachyury,* several *Wnt* family members, *Caudal*, and *Notch*. Therefore, we examined whether cells of clusters 4 and 5 express *Wnt*, *Caudal*, and *Notch.*

A previous report identified eight members of the *Wnt* family in *B. floridae*: *Wnt1*, *Wnt3, Wnt4, Wnt5, Wnt6, Wnt7, Wnt8,* and *Wnt11* (Schubert et al., 2001). Among these, *Wnt8* was first expressed around the blastopore, followed by *Wnt1* (Schubert et al., 2001). Since an improved assembly of the *B. floridae* genome has been published (Simakov et al., 2020), we reexamined *Wnt* members in the *B. floridae* genome and found 12 *Wnt* family members: *Wnt1*, *Wnt2*, *Wnt3, Wnt4, Wnt5, Wnt6, Wnt7, Wnt8, Wnt9, Wnt10, Wnt11*, and *Wnt16* (**Supplementary Figure 7**). Our scRNA-seq analysis demonstrated distinct expression of *Wnt1*, *Wnt4,* and *Wnt8* in *B. japonicum* embryos (**Supplementary Figure 8**), while *Wnt3* and *Wnt7* were expressed at low levels (**Supplementary Figure 8**). A high level of *Wnt1* expression was found in cluster 5 at stages from late gastrula to midneurula (**Figures 4A,C**), and moderate *Wnt8* expression in cluster 4 at stages from late gastrula to midneurula (**Figures 4A,C**). *Wnt4* was expressed in clusters 4 and 5, but the expression level was lower than those of *Wnt1* and *Wnt8* (**Figures 4A,C**). These changes are consistent with previously reported expression patterns of *Wnt* genes (Holland et al., 2000; Schubert et al., 2000). Expression profiles of the three *Wnt* genes in the two clusters, therefore, are not identical. Cluster 5 contains cells with higher expression of *Wnt1* and cluster 4 has higher expression of *Wnt8* (**Figures 4A,C**), suggesting that the two clusters have different properties.

Amphioxus *Caudal* (*Cdx*) was reportedly expressed in the posterior mesendoderm and neural plate of early neurula and in the posterior endoderm, the walls of the neurenteric canal, and the posterior part of the nerve cord in late neurula (Brooke et al., 1998). Expression was also found in embryonic regions including the tail bud, overlapping with that of *Wnt3* (Schubert et al., 2001; Osborne et al., 2009). The present scRNA-seq analysis showed a high level of *Cdx* expression in cells of cluster 5 at stages from late gastrula to early swimming larva, but not in cluster 4 (**Figures 4A,C**; **Supplementary Figure 9A**). Expression of *Wnt3* was detected in cells of cluster 5 (**Supplementary Figure 8C**). These results suggest that cells of cluster 5 occupy the posterior region of the tail bud. In addition to clusters 4 and 5, *Cdx* expression was detected in cells of cluster 0 at late gastrula and cluster 14 at late neurula (**Figure 4**). Future studies might explore *Cdx* expression in these clusters.

On the other hand, *Notch* expression was very broad and found in cells of most clusters from 0 to 13 (**Figure 4**; **Supplementary Figure 9B**). High *Notch* expression was detected in clusters 5, 6, 7, and 13, but only a low level was seen in cluster 4 (**Figure 4**). In summary, the combinatorial expression of *Brachyury*, *Caudal*, *Notch*, and several *Wnt* genes in clusters 4 and 5 provides additional support for the conclusion that these cells comprise the tail bud in amphioxus neurula.

Clusters of Cells With Hox Gene Expression

In contrast to the typical cluster of 13 Hox genes in metazoan genomes (Garcia-Fernandez and Holland, 1994), amphioxus duplicated the posterior Hox genes to 15 genes (Minguillon et al., 2005; **Supplementary Figure 10**). Among them, *Hox1, Hox3*, and *Hox4* are expressed with spatiotemporal collinearity in the developing neural tube of *B. floridae* neurulae (Holland and Garcia-Fernandez, 1996; Wada et al., 1999; Schubert et al., 2004; Pascual-Anaya et al., 2012). *Hox1* is expressed in the developing neural tube of the middle and posterior parts of neurulae, and *Hox1, Hox3*, and *Hox4* show segmental modulation of expression levels, a two-segment phasing of spatial collinearity. Although previous studies focused mainly on *Hox* expression in the developing central nervous system, their results also showed *Hox* expression in the developing mesoderm (somites) as well (Holland and Garcia-Fernandez, 1996; Wada et al., 1999). The present scRNA-seq analysis of *Hox1, Hox3,* and *Hox4* expression profiles demonstrated an additional spatio-expression profile of *Hox* genes, providing new insight to understand evo-devo mechanisms of cephalochordate body plan formation, especially embryonic regions with *Brachyury* expression (**Figure 4** and **Supplementary Figure 11**).

First, in relation to temporal collinearity, initiation of high-level expression of *Hox1, Hox3*, and *Hox4* was detected at early neurula, midneurula, and late neurula, respectively (**Figures 4B,C** and **Supplementary Figure 11**). That is, this result confirmed the previous study of temporal collinearity of gene expression: the most-anterior *Hox1* expressed first, followed by anterior *Hox3*, and then middle *Hox4.* Second, in relation to spatial expression, the present analysis showed that expression profiles of *Hox1*, *Hox3*, and *Hox4* very much resemble that of *Bra2*, although the timing of *Bra2* expression was earlier than that of *Hox1*, *Hox3*, and *Hox4* (**Figures 4B,C**). Specifically, four major clusters with *Hox1, Hox3*, and *Hox4* expression included 4 (tail bud), 5 (tail bud), 8 (somite), and 9 (notochord), in which *Bra2* was specifically expressed (**Figure 4**). For example, in cluster 5, *Hox1* expression was first detected at late gastrula. It was highest at early neurula and decreased at midneurula, whereas *Hox3* expression was first detected at early neurula, was highest at mid- and late neurula, and decreased at early swimming larva (**Figure 4C**). *Hox4* expression was detected at late neurula and decreased at early swimming larva (**Figure 4C**). A similar profile of *Hox* expression was evident in clusters 4, 8, and 9 (**Figures 4A–C**). This suggests that a combined regulatory network in expression and function of *Brachyury* and *Hox* is involved in the formation of somites, notochord, and tail bud of amphioxus embryos.

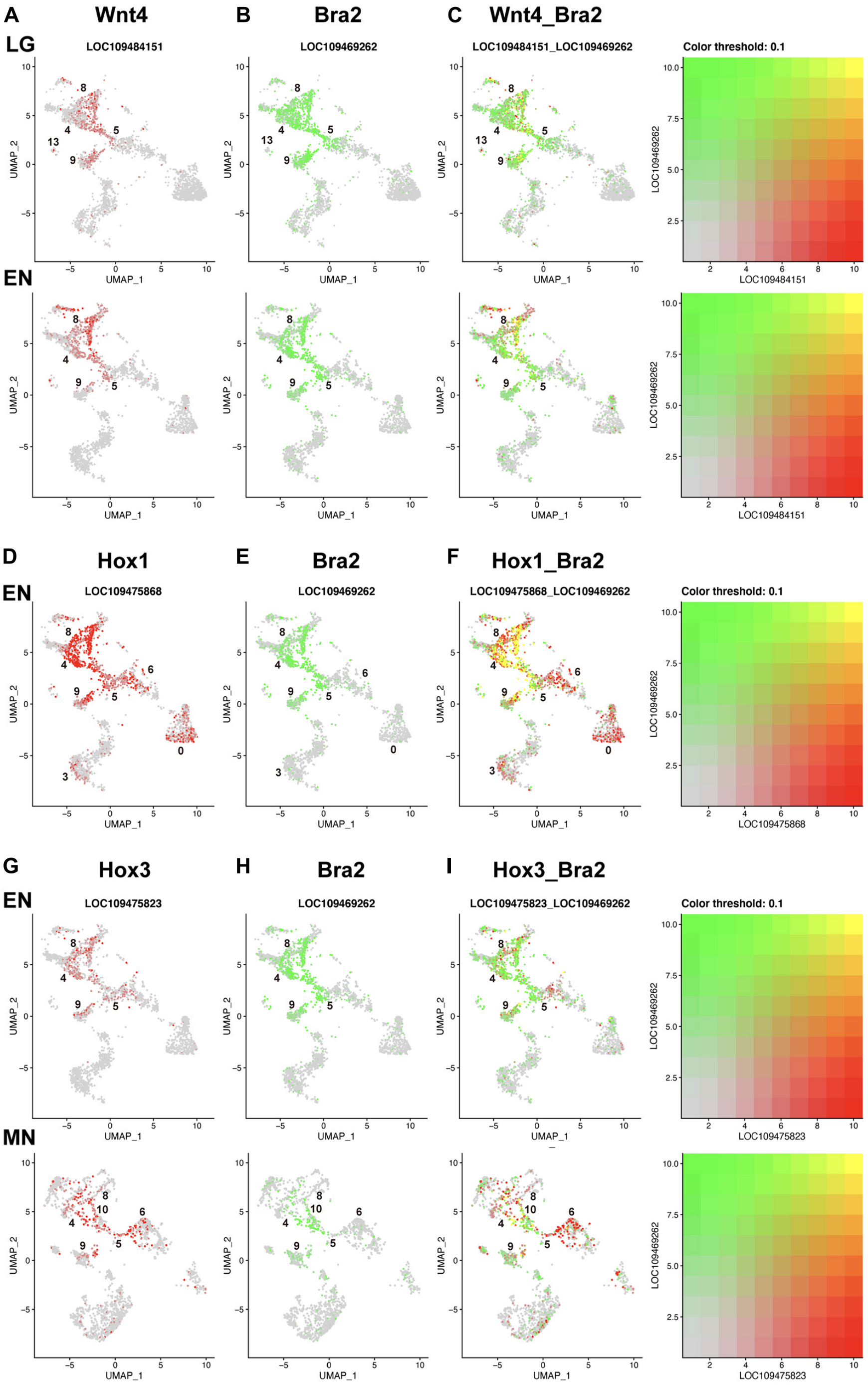


FIGURE 5 | UMAP plots showing co-expression of *Wnt* and *Brachyury*, and *Hox* and *Brachyury* in amphioxus embryos. **(A)** *Wnt4* (red), **(B)** *Bra2* (green), and **(C)** co-expression (yellow) of the two genes at late gastrula (LG) and early neurula (EN) stages. Co-expression is seen in cells of clusters 8, 5, and 9. **(D)** *HoxA1* (red), **(E)** *Bra2* (green), and **(F)** co-expression (yellow) of the two genes at and early neurula (EN) stage. Co-expression is seen in cells of clusters 8, 4, 5, and 9. **(G)** *Hox3* (red), **(H)** *Bra2* (green), and **(I)** co-expression (yellow) of the two genes at early neurula (EN) and midneurula (MN) stages. Co-expression is seen in cells of clusters 8, 4, 5, 9, and 10.

Co-expression of *Bra2* and *Wnt* or *Hox*

As shown in previous sections, clusters 4, 5, 8 and 9, which have *Bra2* expression, also expressed some *Wnt* and *Hox* genes. Therefore, we examined the grade of co-expression of these genes at the single-cell level by superimposing UMAP plots of the two genes (**Figure 5**). For example, *Wnt4* was highly expressed in cluster 8 at late gastrula and early neurula, but not in cells of clusters 4, 5, and 9 (**Figure 4**). Cells of these four clusters were classified to those expressing *Bra2*. Superimposed UMAP plots of the two genes (**Figure 5C**) indicated co-expression of *Wnt4* and *Bra2* in cells of the blastopore, presumptive somite cells, and presumptive notochord cells at late gastrula and early neurula stages.

Compared to *Wnt* genes, *Hox* genes showed greater co-expression with *Brachyury* in early amphioxus embryos (**Figures 5D–F,G–I**). *Hox1* showed higher expression in clusters 4, 5, 8, and 9 at early neurula stage (**Supplementary Figure 11A**), while *Bra2* was expressed in the same clusters at the same stage (**Figure 4B**). High, single-level co-expression of *Hox1* and *Bra2* was found in many cells of clusters 4, 5, 8, and 9 (**Figures 5D–F**). Therefore, cells of the blastopore, presumptive somite cells, and presumptive notochord cells at early neurula stages co-expressed *Hox1* and *Bra2.*

Expression *of Hox3* was found in clusters 4, 5, 8, and 9 at neurula stages (**Figure 4**), and this spatial expression profile was similar to that of *Bra2* (**Figure 4**). Co-expression of *Hox3* and *Bra2* was found in many cells of clusters 4, 5, 8, and 9 (**Figures 5G-I**). Therefore, in addition to *Hox1*, cells of the blastopore, presumptive somite cells, and presumptive notochord cells at early neurula stages co-expressed *Hox3* and *Bra2.*

DISCUSSION

Gene Expression Profiles of scRNA-seq Analyses and Modes of Embryogenesis

scRNA-seq analyses of gene expression profiles have been reported in early embryonic cells of the sea urchin *Strongylocentrotus* (Foster et al., 2020), the ascidian *Ciona* (Sharman et al., 1999; Cao et al., 2019), zebrafish (Wagner et al., 2018), and *Xenopus* (Briggs et al., 2018). In both echinoderms and ascidians, the number of constituent cells is not very large, approximately 2,000–3,000 cells, even at early larval stages. In addition, embryonic cells are likely specified and differentiated into certain types of organs/tissues in restricted lineages. Accordingly, scRNA-seq analysis resulted in comparatively clear-cut clustering of embryonic cells (Sharman et al., 1999; Cao et al., 2019; Foster et al., 2020). In contrast, embryogenesis of zebrafish and *Xenopus* proceeds by a gradual specification pattern, first regionalizing into three germ layers, ectoderm, endoderm, and mesoderm, and later gradually differentiating to tissues and organs. The number of constituent cells exceeds 20,000 cells at the neurula stage. Therefore, the number of clusters classified by the scRNA-seq method is rather small at early stages, but gradually increases as embryogenesis proceeds (Briggs et al., 2018; Wagner et al., 2018). The mode of amphioxus embryogenesis resembles that of vertebrates (Conklin, 1932; Whittaker, 1997; Holland and Holland, 1998). Therefore, our present scRNA-seq analysis resulted in only 10 clusters at midgastrula stage. In addition, constituent cells of amphioxus embryos increase in number after gastrulation, as in the case of vertebrates. Therefore, the cell sample size examined in this study was depauperate, especially in later stages. We are planning to pursue further analyses to sample a sufficient number of *B. japonicum* embryonic cells in the near future. Nevertheless, this scRNA-seq analysis permitted us to survey cells with *Brachyury* expression.

Clustering of Amphioxus Embryonic Cells by scRNA-seq

Using a standard bioinformatic protocol for scRNA-seq analysis, we classified embryonic cells of *B. japonicum* into 15 clusters (0 to 14) (**Figure 2**). Based on results of previous studies that showed specific expression of myogenetic factor genes in the developing paraxial mesoderm or somites, we determined that cluster 8 comprises cells that form somites (**Figure 4**). Then, we showed that *Brachyury* is expressed in cells of clusters 4, 5, 8, and 9 (**Figure 4**). Based on data of other genes, we concluded that cluster 9 belongs to the notochord, and clusters 4 and 5 comprise cells of the blastopore and the developing tail bud. These notions were substantiated by simultaneous expression of various transcription factor genes reportedly expressed in the blastopore and notochord (**Figure 4**). An overall view of clustering by UMAP plots indicates that clusters 4, 5, 8, and 9 have some shared affinity and that they originated from mesoderm (**Figure 2**).

On the other hand, although UMAP plots (**Figure 2**) and gene expression profiles (**Figure 4**) suggest that cluster 0 belongs to ectoderm; clusters 1, 2, 3, 11, and 12 belong to endoderm; and clusters 6, 7, and 10 belong to mesoderm, this study could not unambiguously characterize these clusters, which should be determined in further scRNA-seq analysis. Cluster 13 contains cells with a unique profile of gene expression including *troponin I*, *troponin T*, *gsc*, *hh*, *neterin*, and *Notch* (**Figure 4**). This cluster also occupies a discrete position in UMAP plots (**Figure 2**). Characterization of cluster 13 is therefore an interesting subject of future scRNA-seq analysis.

Development of Embryonic Cells With *Brachyury* Expression

Brachyury was duplicated in cephalochordate genomes, independently of *Brachyury* in other deuterostomes (Holland et al., 1995; Inoue et al., 2017). The present scRNA-seq study showed that *Bra2* is expressed at higher levels and in broader embryonic regions than *Bra1* (**Figure 6**). Since *Bra2*, but not *Bra1*, is expressed in cells of the blastopore (Yuan et al., 2020; this study) and since *Brachyury* expression in the blastopore or archenteron invagination region is a feature shared by most metazoans, *Bra2* is the ancestral *Brachyury* and *Bra1* is a duplicated copy, with more divergent properties, as proposed by previous studies (Tominaga et al., 2018; Yuan et al., 2020). *Bra2* was strongly expressed first in the archenteron invagination region and functions in gastrulation, then in invagination-like movements of cells of the upper archenteron to form the notochord and that of cells of bilateral sides of the archenteron to form somites, and contiguously from the blastopore, in tail bud formation at the posterior-most region of amphioxus embryos. On the other hand, *Bra1* likely retained the potential to form the notochord and somites, serving a supplementary role for *Bra2* in the formation of these organs. These results suggest differences between the two genes in upregulatory

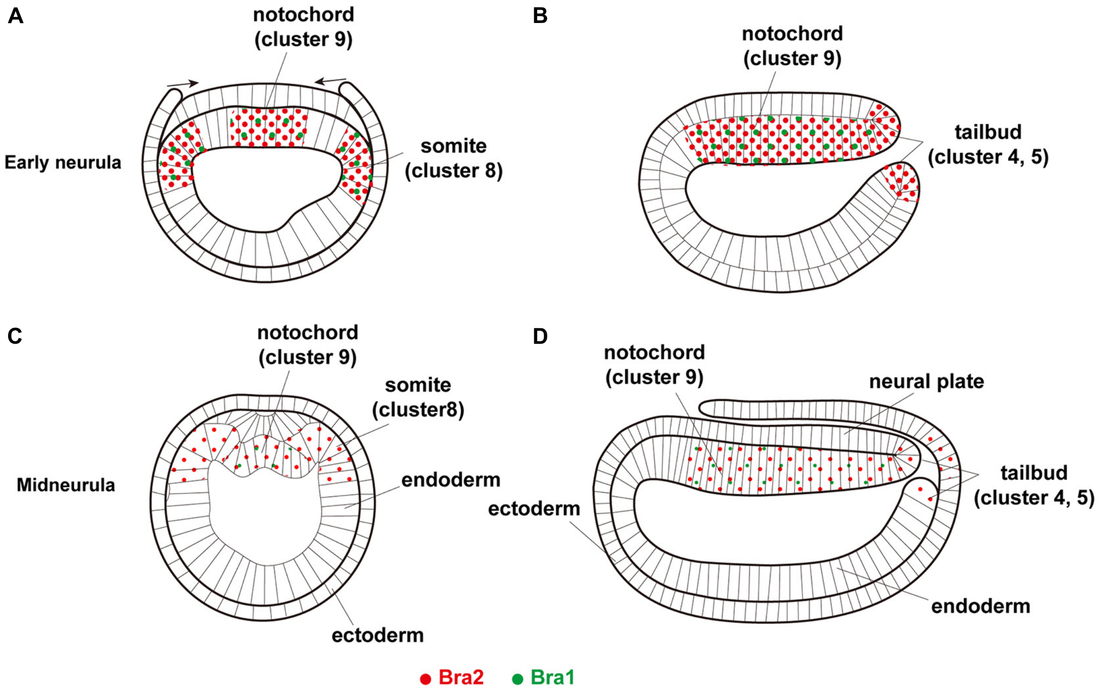


FIGURE 6 | A summary diagram showing expression of two *Brachyury* genes, *Bra1* (green dots) and *Bra2* (red dots), in **(A,B)** early neurula, **(A)** cross section and **(B)** sagittal section, and **(C,D)** midneurula, **(C)** cross section and **(D)** sagittal section. *Bra1* is weakly expressed in the somite (cluster 8) and notochord (cluster 9) at early neurula and becomes undetectable in cluster 8 at midneurula, while a high level of *Bra2* expression is found in the somite (cluster 8), notochord (cluster 9), and tail bud (clusters 4 and 5). Original drawing of embryos by Schubert et al. (2001).

transcriptional control, downregulatory transcriptional control of target genes, and combinatorial transcriptional control of other transcription factor genes in the formation of the blastopore, notochord, somites, and tail bud, which should be examined in future studies.

From an evo-devo point of view, the cephalochordate notochord is unique in having muscle-like properties (Ruppert, 1997). Supporting the results of previous studies (Suzuki and Satoh, 2000; Inoue and Satoh, 2018), the present scRNA-seq analysis demonstrated expression of muscle-structural genes in cells of not only cluster 8 (somites), but also cluster 9 (notochord). In other words, during evolution of cephalochordates, *Brachyury* recruited muscle structural genes under its expressional control. Even more intriguing is the complexity of muscle structural genes. Many muscle structural genes are expressed in the developing somites, while some muscle structural genes are also expressed in the developing notochord, and some are expressed in both somites and notochord (**Figure 4**). Therefore, the muscle-like properties of the amphioxus notochord did not appear by simply copying somites. Future studies should explore molecular and cellular mechanisms involved in this unique muscle gene array of the amphioxus notochord.

The present study also revealed differences in the temporal specification of the notochord and somites. Notochord cluster 9 appeared at the midgastrula stage, suggesting that specification of cells that give rise to the notochord have already occurred (**Figures 2**, **4B**). On the other hand, somite cluster 8 was not formed at the midgastrula stage and first appeared at the late gastrula stage (**Figures 2**, **4B**). Although further studies are required to clarify this issue, these results suggest that the notochord specification occurs earlier than somite specification in amphioxus gastrulae.

Expression of Transcription Factor Genes in Cells With *Brachyury* Expression

Cephalochordates are a research target to understand the origin and evolution of chordates (Holland et al., 2015; Satoh, 2016; Gee, 2018). To date, a large number of evo-devo studies have cloned cDNA of genes for major transcription factors and signaling molecules and extensively examined their spatiotemporal expression profiles by WMISH (Holland et al., 1998; Jackman et al., 2000; Meulemans and Bronner-Fraser, 2002; Schubert et al., 2003; Yu et al., 2007; Wu et al., 2011; Wang et al., 2016). Many such genes show overlapping expression with other genes due to the gradual specification mode of amphioxus embryos. Because *Bra2* is expressed broadly in the embryonic blastopore, somites, notochord, and tail bud, cells with *Bra2* expression simultaneously expressed other transcription factor genes. Here, we discuss only simultaneous expression of *Bra2* and Hox cluster genes, including *Hox1, Hox3*, and *Hox4* in cells of clusters 4, 5, 8, and 9. This was clearly shown in UMAP plot figures (**Figure 5**). The presence of cells with co-expression of *Bra2* and *Hox1, Hox3*, and *Hox4* has not been reported previously. Although details should be explored in future studies, the similarity of expression profiles in cells of clusters 4 (blastopore and tail bud), 5 (blastopore and tail bud), 8 (somites), and 9 (notochord) suggests coordinated roles of *Bra2* and *Hox* in the development of these major organs during amphioxus body plan formation.

Contiguous with the blastopore, the tail bud contains cells with various properties, including cells of the nervous system, mesoderm, and endoderm. As reported in *Xenopus* (Gentsch et al., 2013), *Bra2* expression in the nervous system might be involved in joint regulation of embryonic neuro-mesodermal bipotency in amphibian embryos as well. *Brachyury, Wnt*, genes*, caudal,* and *Notch* are expressed in the tail bud in a coordinated fashion (Holland, 2002). This scRNA-seq analysis showed simultaneous and preferential expression of *Brachyury,* multiple *Wnt* genes, and *caudal* in cells of the tail bud, while *Notch* expression was found in the tail bud, but was rather broadly seen in other embryonic regions as well (**Figure 4**). In addition, this scRNA-seq analysis indicates that clusters 4 and 5 contain cells with simultaneous expression of these genes, suggesting the presence of two different cell types in the developing tail bud. Although we could not distinguish different properties of cells in clusters 4 and 5 at present, this may be an important focus of future studies, because the tail bud is an evolutionary novelty that occurred in chordate embryos and is profoundly involved in the establishment of the chordate-specific body plan.

MATERIALS AND METHODS

Branchiostoma japonicum

Branchiostoma japonicum has been maintained from generation to generation for more than 10 years, first in an aquaculture system at Kumamoto University (Yasui et al., 2013) and then at Tateyama Marine Laboratory of Ochanomizu University in Japan. Adults with mature eggs or sperm were transferred to the Marine Genomics laboratory of Okinawa Institute of Science and Technology Graduate University (OIST). Spawning was induced by controlling temperature and light. Naturally spawned eggs and sperm were mixed to achieve fertilization. All embryos used in this study resulted from mating of one male and one female. Embryos were cultured in Petri dishes filled with filtered natural seawater (pore size, 1 μm). We collected embryos and larvae at six different stages including midgastrula (G3 stage), late gastrula (G5 stage), early neurula (N0 stage), midneurula (N3 stage), late neurula (N5 stage), and early swimming larva (T1 stage). Staging was based on embryonic gross morphology, referring to descriptions of developmental stages of *B. floridae* (Whittaker, 1997) and *B. lanceolatum* (Carvalho et al., 2021).

Embryonic Cell Dissociation and Single-Cell RNA Sequencing

Embryos or larvae at appropriate stages were collected and washed with 1.2 M glycine solution three or four times. For prehatching stages, the fertilization membrane was removed. Then, embryos or larvae were pipetted in 1.2 M glycine solution on ice until cells were completely dissociated. Dissociated cells were counted using hemocytometers (C-chip, NanoEnTek DHC-N01) and diluted with 1.2 M glycine solution to reach an appropriate concentration for the scRNA-seq protocol. Single-cell encapsulation, cDNA synthesis, and library preparation were performed using Chromium Single-Cell 3′ Reagent Kit v3 Chemistry. Libraries were sequenced on an Illumina NovaSeq. Sequencing reads of the six embryonic stages are summarized in **Supplementary Table 1**. Except for midgastrula (350 million reads), more than 600 million reads were obtained for each stage (**Supplementary Table 1**).

Clustering of Cells That Differed in Gene Expression

Single-cell unique molecular identifier (UMI) counting (counting of unique barcodes given to individual transcript molecules) was performed using Cell Ranger Single-Cell Software Suite 4.0.0 from 10 × Genomics. We aligned reads to the *B. belcheri* assembly Haploidv18h27 (GCF_001625305.1) using CellRanger. CellRanger gene expression matrices were further analyzed using the R package Seurat v 4.0.1 (Butler et al., 2018; Stuart et al., 2019; Hao et al., 2020).

Genomes of *B. floridae* (Putnam et al., 2008; Simakov et al., 2020) and *B. belcheri* (Huang et al., 2014) have been decoded, but the *B. japonicum* genome has not been sequenced yet. However, previously, *B. japonicum* was called *B. belcheri,* and taxonomically reclassified as a new species independent of *B. belcheri* approximately 15 years ago (Zhang et al., 2006). Many previous cDNA cloning studies using *B. belcheri*, especially those done by Japanese researchers, likely used this name. Therefore, we here used the *B. belcheri* gene models as references for RNA read annotation.

The number of RNA-seq reads and cells examined at each stage are shown in **Supplementary Tables 1,2**. **Supplementary Table 2** also shows the number of cells examined in each of the 15 clusters. Between 500 and 2500 cells were included in downstream analysis. Individual datasets were normalized by scaling gene expression in each cell by total gene expression, followed by log transformation. The top 2000 highly variable genes across datasets were then used to integrate datasets. Individual time-point datasets were integrated (employing the Seurat v4 pipeline) to identify conserved cell populations across datasets. This technique involves pairwise comparisons of individual cells across multiple datasets followed by hierarchical clustering. UMAP is a dimensional reduction technique. Projection and clustering analysis for visualization of integrated data was conducted using 36 parameter dimensions and a resolution of 0.5 (Stuart et al., 2019). Thirty-six dimensions (i.e., inclusion of 36 principal components) were used in consideration of principal component heatmaps, which show sources of heterogeneity in a dataset, and the ElbowPlot function, which depicts the number of principal components that include variance present in the data. The dataset was also visualized at a resolution of 3 to provide an example of how additional cell states may be revealed, including subtypes of states seen and identified at a resolution 0.5. No one resolution setting is optimal for all clusters, but these disparate settings are intended to assist the reader in data interpretation and identification of candidate genes. The resolution parameter of the FindClusters function can be modulated to show more or fewer clusters and a series of different resolutions can be tested before choosing a value that is appropriate for the biological context of an experiment.

Molecular Phylogeny Analysis

Phylogenetic relationships of family member genes for MFs, T-box, Wnt, and Hox in amphioxus genomes were examined using ORTHOSCOPE (Inoue and Satoh, 2018)[1]. In short, BLAST hit sequences were screened using an *E*-value cutoff of $< 10^{-3}$ and the top five BLAST hits were selected for subsequent analyses. Phylogenetic trees were estimated using the NJ method with the first and second codon positions. To evaluate robustness of internal branches, 100 bootstrap replications were calculated for each dataset.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found at the NCBI BioProject (http://www.ncbi.nlm.nih.gov/bioproject/) under BioProject accession PRJDB10575 and BioSample accessions SAMD00294414–SAMD00294419.

AUTHOR CONTRIBUTIONS

NS, HT, and KN designed the study. MK cultured *Branchiostoma japonicum*. HT and KN carried out molecular lab work. KH and JI analyzed data, carried out statistical analysis, and imaged results. NS wrote the first draft of the manuscript. All authors commented on and approved the final manuscript.

FUNDING

This work was supported by the Okinawa Institute of Science and Technology Graduate School (OIST) to Marine Genomics Unit (NS). This work was partly performed as a project of the Joint Usage/Education Center by the Ministry of Education, Culture, Sports, Science and Technology in Japan.

ACKNOWLEDGMENTS

We thank Jr-Kai Yu for critical reading of the draft manuscript and useful comments for the improvement of the manuscript. Manabu Fujie of the DNA Sequencing Division of OIST is acknowledged for his help in RNA sequencing. Steven Aird (www.sda-technical-editor.org) is also acknowledged for English editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.696875/full#supplementary-material

---

[1] https://www.orthoscope.jp

REFERENCES

Aase-Remedios, M. E., Coll-Llado, C., and Ferrie, D. E. (2020). More than one-to-four via 2R: evidence of an independent amphioxus expansion and two-gene ancestral vertebrate state for MyoD-related myogenic regulatory factors (MRFs). *Mol. Biol. Evol.* 37, 2966–2982. doi: 10.1093/molbev/msaa147

Araki, I., Terazawa, K., and Satoh, N. (1996). Dulication of an amphioxus myogentic bHLH gene is independent of vertebrate myogenic bHLH gene duplication. *Gene* 171, 231–236. doi: 10.1016/0378-1119(96)00174-6

Bertrand, S., and Escriva, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., et al. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* 444, 85–88. doi: 10.1038/nature05241

Briggs, J. A., Weinreb, C., Wagner, D. E., Medason, S., Peshkin, L., Kirschner, M. W., et al. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* 360:eaar5780. doi: 10.1126/science.aar5780

Brooke, N. M., Garcia-Fernàndez, J., and Holland, P. W. H. (1998). The ParaHox gene cluster is anevolutionary sister of the Hox gene cluster. *Nature* 392, 920–922. doi: 10.1038/31933

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. doi: 10.1038/nbt.4096

Cao, C., Lemaire, I. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., et al. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571, 349–354. doi: 10.1038/s41586-019-1385-y

Carvalho, J. E., Lahaye, F., Yong, L. W., Croce, J. C., Escrivá, H., Yu, J.-K., et al. (2021). An updated staging system for cephalochordate development: one table suits them all. *bioRxiv* [Preprint] doi: 10.1101/2020.05.26.112193

Conklin, E. G. (1932). The embryology of amphioxus. *J. Morph.* 54, 69–151. doi: 10.1002/jmor.1050540103

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. doi: 10.1038/nature04336

Foster, S., Oulhen, N., and Wessel, G. (2020). A single cell RNA sequencing resource for early sea urchin development. *Development* 147:dev191528.

Garcia-Fernandez, J., and Holland, P. W. H. (1994). Archetypal organization of the amphixus Hox gene cluster. *Nature* 370, 563–566. doi: 10.1038/370563a0

Gee, H. (2018). *Across the Bridge: Understanding the Origin of the Vertebrates*. Chicago, IL: The University of Chicago Press.

Gentsch, G. E., Owens, N. D. L., Martin, S. R., Piccinelli, P., Faial, T., Trotter, M. W. B., et al. (2013). In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Rep.* 4, 1185–1196. doi: 10.1016/j.celrep.2013.08.012

Gross, J. M., and McClay, D. R. (2001). The role of Brachyury (T) during gastrulation movements in the sea urchin *Lytechinus variegatus*. *Dev. Biol.* 239, 132–147. doi: 10.1006/dbio.2001.0426

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M. III, Zheng, S., Butler, A., et al. (2020). Integrated analysis of multimodal single-cell data. *bioRxiv* [Preprint] doi: 10.1101/2020.10.12.33533

Hirakow, R., and Kajita, N. (1994). Electron miscroscopic study of development of amphixus. *Acta Anat. Nippon.* 69, 1–13.

Holland, L. Z. (2002). Heads or tails? Amphioxus and the evolution of anterior-posterior patterning in deuterostomes. *Dev. Biol.* 241, 209–228. doi: 10.1006/dbio.2001.0503

Holland, L. Z., Holland, N. D., and Schubert, M. (2000). Developmental expression of *AmphiWnt1*, an amphioxus gene in the Wnt1/wingless subfamily. *Dev. Genes Evol.* 10, 522–524.

Holland, L. Z., Venkatesh, T. V., Gorlin, A., Bodmer, R., and Holland, V. D. (1998). Characterization and developmental expression of AmphiNk2.2, an

NK2 class homeobox gene from amphioxus (Phylum Chordata; Subphylum Cephalochordata). *Dev. Genes Evol.* 208, 100–105. doi: 10.1007/s004270050159

Holland, N. D., and Holland, L. (1998). Developmental gene expression in amphioxus: new insight into the evolution. *Am. Zool.* 38, 647–658.

Holland, N. D., Holland, N. D., and Holland, P. W. (2015). Scenarios for the making of vertebrates. *Nature* 520, 450–455. doi: 10.1038/nature14433

Holland, P. W. H., and Garcia-Fernandez, J. (1996). Hox genes and chordate evolution. *Dev. Biol.* 173, 382–395.

Holland, P. W., Koschorz, B., Holland, L. Z., and Herrmann, B. G. (1995). Conservation of Brachyury (T) genes in amphioxus and vertebrates: developmental and evolutionary implications. *Development* 121, 4283–4291. doi: 10.1242/dev.121.12.4283

Huang, S., Chen, Z., Yan, X., Yu, T., Huang, G., Yan, Q., et al. (2014). Decelerated genome evolution in modern vertebrates revealed by analysis of multiple lancelet genomes. *Nat. Commun.* 5:5896. doi: 10.1038/ncomms6896

Imai, K., Levine, M., Satoh, N., and Satou, Y. (2006). Regulatory blueprint for a chordate embryo. *Science* 312, 1183–1187. doi: 10.1126/science.1123404

Inoue, J., and Satoh, N. (2018). Deuterostome genomics: lineage-specific protein expansions that enabled chordate muscle evolution. *Mol. Biol. Evol.* 35, 914–924. doi: 10.1093/molbev/msy002

Inoue, J., Yasuoka, Y., Takahashi, H., and Satoh, N. (2017). The chordate ancestor possessed a single copy of Brachyury gene for notochord acquisition. *Zool. Lett.* 3:4. doi: 10.1186/s40851-017-0064-9

Jackman, W. R., Langeland, J. A., and Kimmel, C. B. (2000). Islet reveals segmentation in the amphioxus hindbrain homolog. *Dev. Biol.* 220, 16–26. doi: 10.1006/dbio.2000.9630

Jiang, D., and Smith, W. C. (2007). Ascidian notochord morphogenesis. *Dev. Dyn.* 236, 1748–1757. doi: 10.1002/dvdy.21184

Lowe, C. J., Clark, D. N., Medeiros, D. M., Rokhsar, D. S., and Gerhart, J. (2015). The deuterostome context of chordate origins. *Nature* 520, 456–465. doi: 10.1038/nature14434

Marletaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene duplication. *Nature* 564, 64–70.

Meulemans, D., and Bronner-Fraser, M. (2002). Amphioxus and lamprey AP-2 genes: implication for neural crest evolution and migration patterns. *Development* 129, 4953–4962. doi: 10.1242/dev.129.21.4953

Minguillon, C., Gardenyes, J., Serra, E., Castro, L. F., Hill-Force, A., Holland, P. W. H., et al. (2005). No more than 14: the endo of amphioxus Hox cluster. *Int. J. Biol. Sci.* 1, 19–23. doi: 10.7150/ijbs.1.19

Munro, E. M., and Odell, G. M. (2002). Polarized basolateral cell motility underlies invagination and convergent extension of the ascidian notochord. *Development* 129, 13–24. doi: 10.1242/dev.129.1.13

Neidert, A. H., Panopoulou, G., and Langeland, J. A. (2008). Amphioxus *goosecoid* and the evolution of the head organizer and prechordal plate. *Evol. Dev.* 2, 303–310. doi: 10.1046/j.1525-142x.2000.00073.x

Oliveri, P., Walton, K. D., Davidson, E. H., and McClay, D. R. (2006). Repression of mesodermal fate by foxa, a key endoderm regulator of the sea urchin embryo. *Development* 133, 4173–4181. doi: 10.1242/dev.02577

Onai, T., Aramaki, T., Inomata, H., Hirai, T., and Kuratani, S. (2015). Ancestral mesodermal reorganization and evolution of the vertebrate head. *Zool. Lett.* 1:29. doi: 10.1186/s40851-015-0030-3

Ono, H., Koop, D., and Holland, L. Z. (2018). Nodal and Hedgehog synergize in gill slit formation during development of the cephalochordate *Branchiostoma floridae*. *Development* 145:dev162586. doi: 10.1242/dev.162586

Osborne, P. W., Benoit, G., Laudet, V., Schubert, M., and Ferrier, D. E. K. (2009). Differential regulation of ParaHox genes by retinoic acid in the invertebrate chordate amphioxus (*Branchiostoma floridae*). *Dev. Biol.* 327, 252–262. doi: 10.1016/j.ydbio.2008.11.027

Pascual-Anaya, J., Adachi, N., Alvarez, S., Kuratani, S., D'Aniello, S., and Garcia-Fernàndez, J. (2012). Broken colinearilty of the amphioxius Hox cluster. *EvoDevo* 3:28. doi: 10.1186/2041-9139-3-28

Petillon, Y. L., Luxardi, G., Scerbo, P., Cibois, M., Leon, A., Subirana, L., et al. (2017). Nodal-activin pathway is a conserved neural induction signal in chordates. *Nat. Ecol. Evol.* 1, 1192–1200. doi: 10.1038/s41559-017-0226-3

Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071.

Ruppert, E. E. (1997). "Cephalochordata (Acrania)," in *Microscopic Anatomy of Invertebrates*, ed. F. W. Harrison (New York, NY: Willey-Liss).

Satoh, G., Takeuchi, J. K., Yasui, K., Tagawa, K., Saiga, H., Zhang, P., et al. (2002). *Amphi-Eomes/Tbr1*: an amphioxus cognate of vertebrate *Eomesodermi*n and *Tbrain1* genes whose expression reveals evolutionarily distinct domain in amphioxus development. *J. Exp. Zool.* 294, 136–145. doi: 10.1002/jez.10149

Satoh, N. (2016). *Chordate Origins and Evolution: The Molecular Evolutionary Road to Vertebrates*. San Diego, CA: Academic Press.

Satoh, N., Tagawa, K., and Takahashi, H. (2012). How was the notochord born? *Evol. Dev.* 14, 56–75. doi: 10.1111/j.1525-142x.2011.00522.x

Schubert, M., Holland, L. Z., Panopoulou, G. D., Lehrach, H., and Holland, N. D. (2000). Characterization of amphioxus *AmphiWnt8*: insights into the evolution of patterning of the embryonic dorsoventral axis. *Evol. Dev.* 2, 85–92. doi: 10.1046/j.1525-142x.2000.00047.x

Schubert, M., Holland, L. Z., Stokes, M. D., and Holland, N. D. (2001). Three amphioxus Wnt genes (AmphiWnt3, AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Dev. Biol.* 240, 262–273. doi: 10.1006/dbio.2001.0460

Schubert, M., Holland, N. D., Escriva, H., Holland, L. Z., and Laudet, V. (2004). Retinoic acid influences anteroposterior positioning of epidermal sensory neurons and their gene expression in a developing chordate (amphioxus). *Proc. Natl. Acad. Sci. U.S.A.* 101, 10320–10325. doi: 10.1073/pnas.0403216101

Schubert, M., Muelemans, D., Bronner-Fraser, M., Holland, L. Z., and Holland, N. D. (2003). Differential mesodermal expression of two *MyoD* family members (*AmphiMRF1* and *AmphiMRF2*). *Gene Expr. Patterns* 3, 199–202. doi: 10.1016/s1567-133x(02)00099-6

Sharma, S., Wang, W., and Stolfi, A. (2019). Single-cell transcriptome profiling of the *Ciona* larval brain. *Dev. Biol.* 448, 226–236. doi: 10.1016/j.ydbio.2018.09.023

Sharman, A. C., Shimeld, S., and Holland, P. W. H. (1999). An amphioxus Msx gene expressed predominantly in the dorsal neural tube. *Dev. Genes Evol.* 209, 260–263. doi: 10.1007/s004270050251

Shimeld, S. M. (1997). Characterization of amphioxus HNF-3 genes: conserved expression in the notochord and floor plate. *Dev. Biol.* 183, 74–85. doi: 10.1006/dbio.1996.8481

Shimeld, S. M. (1999). The evolution of the hedgehog gene family in chordates: insights from amphioxus hedgehog. *Dev. Genes Evol.* 209, 40–47. doi: 10.1007/s004270050225

Shimeld, S. M. (2000). An amphioxus netrin gene is expressed in midline structres during embryonic and larval development. *Dev. Genes Evol.* 210, 337–344. doi: 10.1007/s004270000073

Simakov, O., Marlétaz, F., Yue, J. X., O'Connell, B., Jenkins, J., Brandt, A., et al. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nat. Ecol. Evol.* 4, 820–830. doi: 10.1038/s41559-020-1156-z

Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. *Nat. Rev. Genet.* 20, 257–272.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M. III, et al. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.e21.

Suzuki, M. M., and Satoh, N. (2000). Genes expressed in the amphioxus notochord revealed by EST analysis. *Dev. Biol.* 224, 168–177. doi: 10.1006/dbio.2000.9796

Tagawa, K., Humphreys, T., and Satoh, N. (1998). Novel pattern of *Brachyury* gene expression in hemichordate embryos. *Mech. Dev.* 75, 139–143. doi: 10.1016/s0925-4773(98)00078-1

Tan, X., Zhang, P. J., and Du, S. J. (2014). Evolutionary aspects of a new MyoD gene in amphioxus (*Branchiostoma belcheri*) and its promoter specificity in skeletal and cardiac muscle. *Biologia* 69, 1210–1224. doi: 10.2478/s11756-014-0427-z

Terazawa, K., and Satoh, N. (1997). Formation of the chordamesoderm in the amphioxus embryo: analysis with *Brachyur*y and *fork head/HNF-3* genes. *Dev. Genes Evol.* 207, 1–11. doi: 10.1007/s004270050086

Tominaga, H., Satoh, N., Ueno, N., and Takahashi, H. (2018). Enhancer activities of amphioxus Brachyury genes in embryos of the ascidian, *Ciona intestinalis*. *Genesis* 58:e23240. doi: 10.1002/dvg.23240

Urano, A., Suzuki, M., Zhang, P., Satoh, N., and Satoh, K. (2003). Expression of muscle-related genes and two MyoD genes during amphioxus notochord development. *Evol. Dev.* 5, 447–458. doi: 10.1046/j.1525-142x.2003.03051.x

Wada, Y., Garcia-Fernandez, J., and Holland, P. W. H. (1999). Colinear and segmental expression of amphioxus Hox genes. *Dev. Biol.* 213, 131–141. doi: 10.1006/dbio.1999.9369

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., and Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* 360, 981–987. doi: 10.1126/science.aar4362

Wang, J., Li, G., Qian, G.-H., and Wang, Y.-Q. (2016). Expression analysis of eight amphioxus genes involved in the Wnt/beta-catenin signaling pathway. *Zool. Res.* 37, 136–143.

Whittaker, J. R. (1997). "Cephalochordates, the lancelets," in *Embryology: Constructing the Organism*, eds S. F. Gilbert and A. M. Raunio (Sunderland, MA: Sinauer).

Wu, H.-R., Chen, Y.-T., Su, Y.-H., Luo, Y. J., Holland, L. Z., and Yu, J.-K. (2011). Asymmetric localization of germline markers Vasa and Nanos during early development in the amphioxus *Branchiostoma floridae*. *Dev. Biol.* 353, 147–159. doi: 10.1016/j.ydbio.2011.02.014

Yasui, K., Igawa, T., Kaji, T., and Henmi, Y. (2013). Stable aquaculture of the Japanese lancelet *Branchiostoma* japonicum for 7 years. *J. Exp. Zool. B Mol. Dev. Evol.* 320B, 538–547. doi: 10.1002/jez.b.22540

Yu, J.-K., Mazet, F., Chen, Y.-T., Haung, S.-W., Jung, K.-C., and Shimeld, S. M. (2008). The *Fox* genes of *Branchiostoma floridae*. *Dev. Genes Evol.* 218, 629–638.

Yu, J.-K., Satou, Y., Holland, N. D., Shin-I, T., Kohara, Y., Satoh, N., et al. (2007). Axial patterning in cephalochordates and the evolution of the organizer. *Nature* 445, 613–617. doi: 10.1038/nature05472

Yuan, L., Wang, Y., and Li, G. (2020). Differential expression pattern of two Brachyury genes in amphioxus embryos. *Gene Expr. Patterns* 38:119152. doi: 10.1016/j.gep.2020.119152

Zhang, Q.-J., Zhong, J., Fang, S.-H., and Wang, Y.-Q. (2006). *Branchiostoma japonicum* and *B. belcheri* are distinct lancelets (Cephalochordata) in Xiamen waters in China. *Zool. Sci.* 23, 573–579. doi: 10.2108/zsj.23.573

Zhu, X., Shi, C., Zhong, Y., Liu, X., Yan, Q., Wu, X., et al. (2020). Cilia-driven asymmetric Hedgehog signalling determines the amphioxus left-right axis by controlling Dand5 expression. *Development* 147:dev182469. doi: 10.1242/dev.182469

Conflict of Interest: 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 © 2021 Satoh, Tominaga, Kiyomoto, Hisata, Inoue and Nishitsuji. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 July 2021
doi: 10.3389/fcell.2021.702986

Optical Clearing and Light Sheet Microscopy Imaging of Amphioxus

Simona Machacova[1], Helena Chmelova[2], Anna Vavrova[1], Zbynek Kozmik[1] and Iryna Kozmikova[1]*

[1] Laboratory of Transcriptional Regulation, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czechia, [2] Light Microscopy Core Facility, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czechia

Cephalochordates (amphioxi or lancelets) are representatives of the most basally divergent group of the chordate phylum. Studies of amphioxus development and anatomy hence provide a key insight into vertebrate evolution. More widespread use of amphioxus in the evo–devo field would be greatly facilitated by expanding the methodological toolbox available in this model system. For example, evo–devo research on amphioxus requires deep understanding of animal anatomy. Although conventional confocal microscopy can visualize transparent amphioxus embryos and early larvae, the imaging of later developmental stages is problematic because of the size and opaqueness of the animal. Here, we show that light sheet microscopy combined with tissue clearing methods enables exploration of large amphioxus specimens while keeping the surface and the internal structures intact. We took advantage of the phenomenon of autofluorescence of amphioxus larva to highlight anatomical details. In order to investigate molecular markers at the single-cell level, we performed antibody-based immunodetection of melanopsin and acetylated-α-tubulin to label rhabdomeric photoreceptors and the neuronal scaffold. Our approach that combines light sheet microscopy with the clearing protocol, autofluorescence properties of amphioxus, and antibody immunodetection allows visualizing anatomical structures and even individual cells in the 3D space of the entire animal body.

Keywords: amphioxus, light sheet microscopy, clearing technique, whole mount immunohistochemistry, photoreceptor, acetylated tubulin, melanopsin

OPEN ACCESS

Edited by:
Stephanie BERTRAND,
UMR 7232 Biologie Intégrative des Organismes Marins (BIOM), France

Reviewed by:
Nicholas Holland,
University of California, San Diego, United States
Ricard Albalat,
University of Barcelona, Spain

****Correspondence:***
Iryna Kozmikova
kozmikova@img.cas.cz

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 April 2021*
Accepted: *21 June 2021*
Published: *26 July 2021*

Citation:
Machacova S, Chmelova H, Vavrova A, Kozmik Z and Kozmikova I (2021) Optical Clearing and Light Sheet Microscopy Imaging of Amphioxus. Front. Cell Dev. Biol. 9:702986. doi: 10.3389/fcell.2021.702986

INTRODUCTION

Cephalochordata (also called lancelets or amphioxi) represent the earliest chordate lineage (Delsuc et al., 2006) and share with vertebrates a similar body plan. Therefore, amphioxus is the best animal model to elucidate the origin and evolution of chordates. The studies that investigate the early development of amphioxus embryos prevail over the ones examining the ontogenesis from late larvae to adult animals. Many techniques for imaging early embryos give poor results when applied to late larvae. And yet, asymmetric amphioxus larvae possess plenty of remarkable characteristics such as a preoral pit, an extended oral opening located on the left side, a club-shaped gland, an endostyle, an unpaired gill slits on the right side (while posterior gill slits rotate to medioventral position in early larvae), a Hatschek's groove, a Hatschek's nephridium, a Räder organ (a wheel organ), and an anus moving from the originally right position to medioventral position and then to the left (Hatschek, 1881; Willey, 1891, 1894; Conklin, 1932; Stokes and Holland, 1995). The specimens at later developmental stages exhibit high-density tissue and lose optical transparency typical of amphioxus embryonic and early larval stages. One-month-old larva is already partially opaque, and such techniques as whole-mount *in situ* hybridization or immunohistochemistry followed by confocal microscopy imaging become challenging. In the past, large-specimen analysis and visualization was mainly performed by using mechanical sectioning and imaging of individual slices. Nowadays, single plane illumination microscopy (SPIM; or light sheet microscopy) combined with tissue clearing allows significantly faster examination of a large specimen while retaining the key context of the surrounding tissue (Huisken and Stainier, 2009; Masson et al., 2015; Nie et al., 2020). The desired transparency is achieved by elimination of light scattering on the interface between environments with various refractive indexes, mostly lipid deposits and light-absorbing substances. CUBIC is a hydrophilic clearing method that enables effective tissue clearing in two steps (Susaki et al., 2015). First, CUBIC1 reagent washes out lipids, and second, CUBIC2 substitutes the space instead of missing lipids to equilibrate the refractive index of the tissue. The advantages of the method are safety, usage of commonly available chemicals, and simplicity. In addition, the CUBIC method allows combining tissue clearing with immunohistochemistry and endogenous fluorescence detection.

In this study, we applied the CUBIC clearing method followed by light sheet microscopy imaging to examine 1–, 3–, and 6-month-old *Branchiostoma floridae* specimens. It has been reported that amphioxus possesses its own green fluorescent proteins (GFPs) ubiquitously distributed in larvae. In metamorphic juvenile, GFPs are located in the anterior body portion, predominantly in the support cells of oral cirri, more diffusely in the epidermis (mainly in the anterior and posterior ends of the animal) (Deheyn et al., 2007; Bomati et al., 2009). We analyzed the autofluorescence of larvae and juvenile specimens exposed to four excitation wavelengths (405, 488, 561, and 638 nm). In addition, we performed immunohistochemistry (IHC) staining for acetylated tubulin to visualize the neural net and for melanopsin to label rhabdomeric photoreceptor cells (Koyanagi et al., 2005; Pergner and Kozmik, 2017).

MATERIALS AND EQUIPMENT

Animal Culture, Spawning Induction, and Fixation

Equipment

- Room with controlled temperature and light/dark cycle
- Sea water (Bremerhaven, Alfred-Wegener-Institut)
- Saltwater tank with mechanical filtration and UV sterilizer (AQUA SCHWARZ)
- 5 L tank
- Algae (*Isochrysis lutea*, *Phaeodactylum tricornutum*)
- Plastic cups
- Red-light flashlight
- Stereomicroscope
- 5 and 15 cm plastic Petri dishes coated with 1% agarose to prevent sticking the eggs to the bottom
- Manual centrifuge (∼3,000 rpm)
- 15 ml centrifuge tubes
- Rotator-nutator shaker

Stock solutions

- 16% paraformaldehyde aqueous solution, EM grade, 10 ml ampoule
- MOPS solution [0.1 M 3-(N-morpholino)propanesulfonic acid, 2 mM $MgSO_4$, 1 mM EGTA, 0.5 M NaCl], pH 7.5, storage at −20 °C

Working solution

- Filtered sea water mixed with distilled water (7:1)
- 4% MOPS–PFA, pH 7.5
- 100% methanol

Whole-Mount Immunohistochemistry, Clearing, and Mounting

Equipment

- 10 ml Erlenmeyer flask (for clearing the agarose column)
- Modeling clay
- Glass capillary for light sheet microscopy (size 2, 3)

Stock solutions

- 100% methanol
- 10× TBS buffer (Tris base 200 mM, NaCl 1.5 mM, pH 7.6)
- 10% Triton X-100
- 10% bovine serum albumin (BSA)
- 100% normal donkey serum (NDS)
- Mouse anti-melanopsin primary antibody, generated in our laboratory (Bozzo et al., 2017)
- Mouse anti-melanopsin primary antibody, generated by Koyanagi et al. (2005)
- Rabbit anti-acetyl-α-tubulin antibody (Cell Signaling #5335, Danvers, MA, United States)
- Donkey anti-mouse Alexa 647 (Thermo Fisher A-31571, Waltham, MA, United States)
- Donkey anti-rabbit Alexa 488 (Life Technologies A-21206, Carlsbad, CA, United States)
- 4′,6-Diamidine-2′-phenylindole dihydrochloride (DAPI)
- Urea
- N,N,N′,N′-Tetrakis(2-hydroxypropyl)ethylenediamine (4NTEA)
- Triethanolamine (TEA)
- Sucrose
- Triton X-100

Working solutions

- Methanol series [70, 50, 25% (wt/wt)]
- 1× TBT buffer (1 × TBS, 0.1% Triton X-100), filtered
- Blocking solution [4% (wt/wt) BSA, 40% (wt/wt) NDS in 1× TBT]
- 1.7% low gelling temperature agarose dissolved in distilled water

- CUBIC1—clearing solution [35% (wt/wt) dH_2O, 25% (wt/wt) urea, 25% (wt/wt) 4NTEA, 15% (wt/wt) Triton X-100]; add in order and allow each reagent to dissolve; stir over low heat (~42 °C)
- CUBIC2—refractive index matching solution [15% (wt/wt) dH_2O, 25% (wt/wt) urea, 10% (wt/wt) TEA, 50% (wt/wt) sucrose, 0,1% (wt/wt) Triton X-100]; add in order and allow each reagent to dissolve; stir over low heat; final refractive index 1.473; store for 1–2 months at RT; strong smell of ammonia indicates expired solution

Light Sheet Microscopy Imaging

- Zeiss Lightsheet Z.1 microscope (Carl Zeiss AG, Oberkochen, Germany)
- 10×/0.2 illumination objective
- Detection objective Clr Plan-Neofluar 20×/1.0 Corr nd = 1.45 suited for CUBIC2 clearing media
- Camera PCO.edge 5.5 (sCMOS), 6.5 μm pixel
- Storage for data imaging (~160 TB)
- Stereomicroscope for sample preparation

Software requirements

- ZEN 3.1 (Carl Zeiss AG, black edition)
- Imaris Stitcher ×64 9.6.0 software (Bitplane, Zurich, Switzerland)
- Imaris ×64 9.6.0 software (Bitplane)
- Adobe Illustrator CS6
- Fiji ImageJ

Methods

Animal Culture and Fixation

Adults of *B. floridae* were collected in Old Tampa Bay, Florida, during the spawning season and were cultured at the Institute of Molecular Genetics of the Czech Academy of Sciences (Prague, Czechia). The maintenance of amphioxus culture and spawning induction was described previously (Pergner et al., 2020). Embryos and larvae were raised at 25 °C. Initiation of feeding started 2 days post fertilization with *Isochrysis lutea* and *Phaeodactylum tricornutum*. The animals were fed daily. The larvae were kept in a 90-mm tissue culture dish for 1 month in the dark without day/night cycle. After 1 month, the animals were kept in a 14/10-h day/night cycle. Two-month-old juveniles were transferred to boxes with a volume of 1.5 L and fed daily with *Isochrysis lutea.* Larvae and juveniles in desired stages of development were fixed with 4% MOPS–PFA for 15 min on ice and transferred into 100% methanol, followed by five washes in 100% methanol, 20 min each. Samples were stored at –20 °C.

Tissue Clearing and Immunohistochemistry Staining

(1) Place samples to a four-well plate and transfer them from methanol to TBT through a series of washings with TBT containing 70, 50, and 25% of methanol (for 15 min each, volume 0.5 ml, agitating).
(2) Wash three times in TBT for 15 min (in 0.5 ml, agitating).
(3) With continued agitation, add a few drops of CUBIC1 solution every 30 min. When the well is full, replace the liquid with fresh CUBIC1. Leave the specimens in CUBIC1 on the rotating shaker at room temperature overnight or longer until the total tissue is cleared (depending on the specimen size, it may last for 12–48 h). Replace CUBIC1 by a fresh one every 24 h.

Attention: Fast direct transfer of larger samples into CUBIC1 causes irreversible bending of individuals.

(4) Remove half of the liquid and add the same volume of TBT, then wash the cleared samples for 30 min with agitation.
(5) Wash the samples in TBT alone five times for 20 min with agitation.

(*Optional IHC steps follows*)

(6) Block the specimens dedicated to IHC staining in blocking solution for at least 1 h at room temperature with agitation.
(7) Incubate the samples in TBT with 10% BSA containing primary antibodies overnight at 4 °C with agitation.

In this study, anti-acetyl-α-tubulin (1:500) and anti-melanopsin (1:250) were used. Mouse polyclonal antibody was generated against amphioxus melanopsin as described by Bozzo et al. (2017) using a peptide derived from the C-terminus of *B. floridae* protein.

(8) The next day, wash the samples five times in TBT for 20 min with agitation.
(9) Incubate samples in TBT containing 10% BSA with secondary antibodies for 3 h at room temperature with agitation.

In this study, donkey anti-mouse Alexa 647 (1:500) and donkey anti-rabbit Alexa 488 (1:500) antibodies were used.

(10) Wash the samples in TBT five times for 20 min with agitation.

(*End of IHC steps*)

(11) Stain nuclei with DAPI (1:1,000) in TBT overnight at 4 °C with agitation.
(12) Wash the samples in TBT three times for 20 min at room temperature with agitation.
(13) Transfer the specimens into CUBIC2 solution in the same way as in the case of CUBIC1.
(14) Samples can be stored in CUBIC2 at room temperature until imaging.

Attention: Storage longer than 1 month causes crystallization of CUBIC2, which can be turned back to the liquid state by adding TBT and gentle mixing.

(15) For imaging, immerse the samples into warm agarose (about 42 °C) and suck them into a glass capillary. Choose the size of the glass capillary corresponding to the size of the specimen.
(16) To adjust the agarose refractive index, it has to be incubated in CUBIC2 as well. After agarose polymerization, transfer the capillary into a 10-ml Erlenmeyer flask containing CUBIC2 solution and attach the capillary to the neck of the flask by modeling clay. Gently push out the agarose

column containing specimens into CUBIC2 solution. Left the samples in a dark room for at least 24 h.

Microscopy Imaging and Data Processing

Samples were imaged under a Zeiss Lightsheet Z.1 microscope (Carl Zeiss AG) using the software ZEN 3.1 (black edition, Carl Zeiss AG) and equipped with two 10×/0.2 illumination objectives (for dual-sided illumination), detection objective Clr Plan-Neofluar 20×/1.0 Corr nd = 1.45 suited for CUBIC2 clearing media and two cameras pco.edge 5.5 (sCMOS). Online dual side fusion in ZEN 3.1 (black edition) was used to fuse left and right illuminated images. Individual channels were imaged sequentially in a switch mode between z-stacks (i.e., first, the z-stack image of one fluorophore channel is completed, then it switches to the other). The excitation and emission wavelengths were used in individual channels as follows: blue (excitation 405 nm, emission 420–470 nm), green (excitation 488 nm, emission 505–545 nm), red (excitation 561 nm, emission 575–615 nm), and magenta (excitation 638 nm, emission LP 660 nm).

The specimens were imaged according to the conditions described in **Supplementary Table 1**. The obtained tiles (squares in a regular grid subdividing the image) were stitched in Imaris Stitcher ×64 9.6.0 software (Bitplane). Visualization of the final images was done in Imaris ×64 9.6.0 software version (Bitplane) and Fiji ImageJ software. Postprocessing of the data was done on specialized analysis computers (**Supplementary Table 2**). Final panel figures were assembled in Adobe Illustrator CS6.

RESULTS

Tissue Clearing of Amphioxus

The embryos of *B. floridae* are transparent during the first weeks of development. As the larva grows, it loses its transparency and examination of the whole-mount morphology becomes challenging. To overcome the low transparency of older samples, we applied the hydrophilic tissue clearing method CUBIC. Five individuals of 1-month-old larva along with 3– and 6-month-old juveniles were selected for the following experiments: solely CUBIC clearing or combined with IHC. The intact and cleared larvae and juveniles were observed under a conventional light stereomicroscope (**Figure 1**). The clearing procedure remarkably increased the transparency of the animal specimens. The pigmented cells forming the dorsal ocelli or frontal eye became distinctly visualized in the background of cleared animals.

Autofluorescence and Acetylated Tubulin Staining of 1-Month-Old Larvae

Endogenous fluorescence constitutes both an advantage and an obstacle for imaging procedures. On the one hand, autofluorescence may serve as an endogenous marker of certain structures or cells. On the other hand, endogenous fluorescence could interfere with the desired staining in a given stage of development. Therefore, we explored the amphioxus endogenous fluorescent signal of cleared larvae and juveniles. One-month-old larva with seven unpaired gill slits exhibited a low level of autofluorescence excited by 405, 488, 561, and 638 nm laser wavelengths. In all spectral channels, the fluorescent intensity was the strongest in the gut (**Figures 2A–E**) and frontal eye (**Figures 2B,D,E**). Blue autofluorescence was detected almost homogeneously throughout the tissue (**Figure 2A**). Emission induced by 488 and 638 nm laser was predominantly present in the most dorsal tissue (fin rays lying in the fin chambers), gut, and branchial apparatus (**Figures 2B,D**). Red autofluorescence was the strongest in the gut (**Figure 2C**). Next, we performed combined IHC staining of acetyl-α-tubulin with DAPI (**Figures 2F–Q** and **Supplementary Videos 1, 2**). The laser power required to excite DAPI or Alexa 488 was three times lower than in the case of endogenous fluorescence. Thus, the autofluorescence could not hinder the imaging of exogenous markers. Probably due to the higher density of nuclei, DAPI staining showed stronger intensity in the neural cord with brain vesicle, neuropore oral opening and branchial apparatus, endostyle, and gut, including the ilio-colon ring (**Figures 2F,H,I,K**). Staining with anti-acetyl-α-tubulin antibody clearly visualized the preoral pit, endostyle, gill slits, frontal eye, and nervous system throughout the whole specimen (**Figures 2G,H,J–Q** and **Supplementary Videos 1, 2**).

The regular dark regions in **Figures 2A–K** represent a processing artifact caused by stitching the tiles into one image because of lower light intensity in the camera chip margins. This effect could be avoided by using only the central part of the chip. In this case, the number of tiles would have to be doubled and the imaging time of the whole sample would be increased. The need of stitching can be avoided by usage of real-time adjustable tiling light sheet selective plane illumination microscopy (TLS-SPIM) or multiangle-resolved subvoxel selective plane illumination microscope (Mars-SPIM) (Fu et al., 2016; Nie et al., 2020).

Autofluorescence and Acetylated Tubulin Staining of a 3-Month-Old Juvenile

Furthermore, we tested a 3-month-old juvenile after metamorphosis with 28 pairs of gill slits and roughly 8 mm in length (**Figure 3**). Three-month-old juveniles exhibit right-side position of hepatic diverticulum and medioventral position of both atriopore and anal opening. In general, the detected autofluorescence had similarly low intensity as in the case of 1-month-old larvae except for the appearance of signal in the oral apparatus (compare **Figures 2A–E**, **3B–F**). Autofluorescence emitted in the blue spectrum was almost undetectable. Therefore, DAPI staining was performed to get the outline of the whole-mount specimen (**Figure 3A**). The DAPI signal was found in all tissues and the strongest intensity was detected in the neural tube, velum, gill slits, and gut. Endogenous GFPs of amphioxus were strongly expressed in oral cirri (**Figures 3B,E,F**). A lower green signal was also detected in the frontal eye. The frontal eye, basis of the oral apparatus, neural tube, hepatic diverticulum, and gut were noticeable in the red spectrum (**Figure 3C**). The autofluorescence signal in the far red spectrum was detected in the frontal eye, hepatic diverticulum, and gut (**Figures 3D–F**). Immunohistochemistry staining for acetyl-α-tubulin revealed advanced progressive branching of neuronal axons when compared with younger

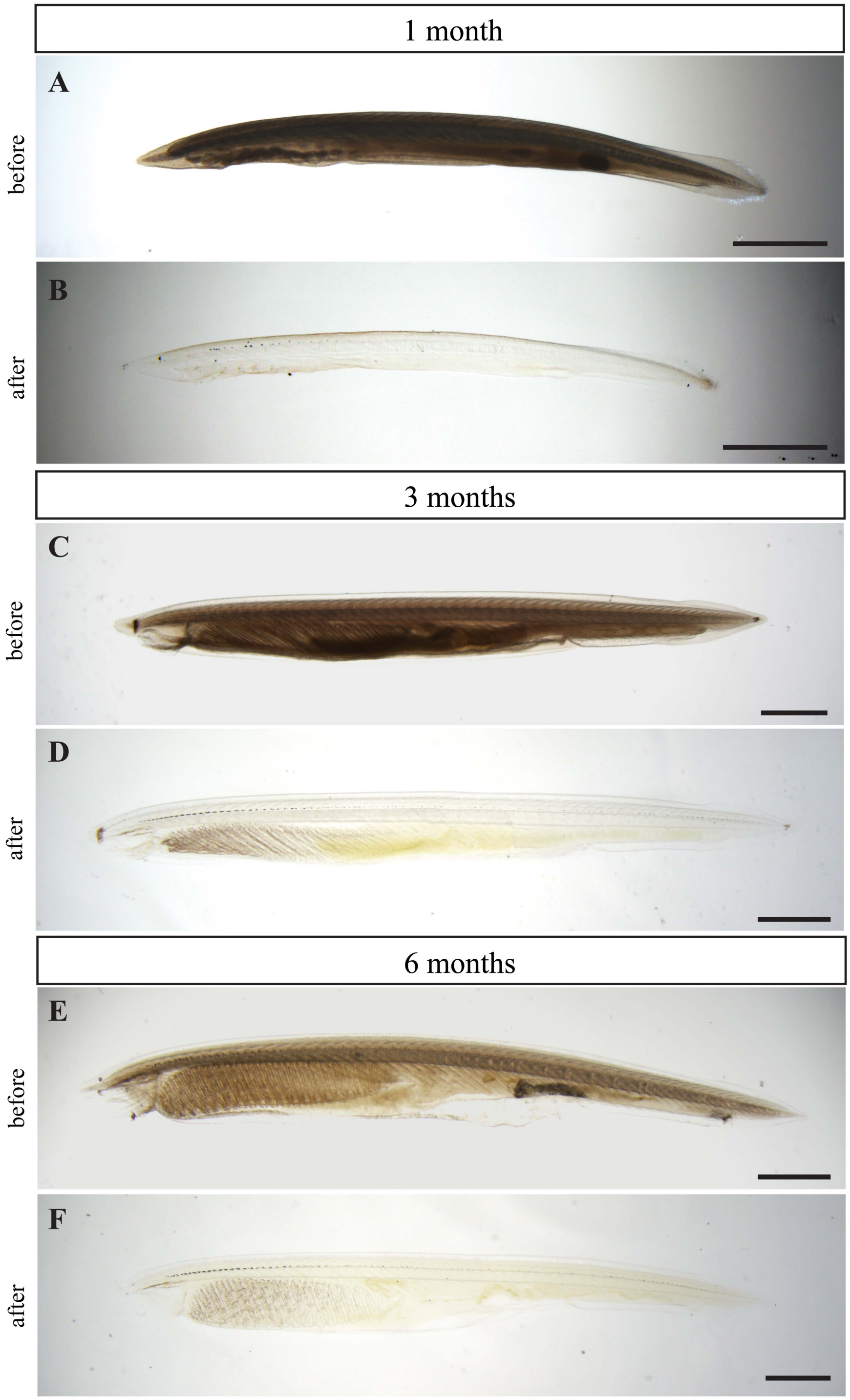


FIGURE 1 | Clearing effect of CUBIC technique on larvae and juveniles of *B. floridae*. One-month-old larva before **(A)** and after clearing **(B)**, 3-month-old juvenile before **(C)** and after clearing **(D)**, and 6-month-old juvenile before **(E)** and after clearing **(F)**. Larvae and juveniles depicted before and after clearing came from the same batch and were raised in the same conditions. Wide field microscopy with transmitted light was used for imaging all specimens. Scale bar 0.5 mm.

larvae (compare **Figures 2F–Q** with **Figures 3G–Q** and **Supplementary Video 3**). Acetylated tubulin marked the neural net with axons extending from the neural tube in dorsal and lateroventral directions. Strong innervation was present in the rostrum, frontal eye, velum, hepatic diverticulum, nerve net surrounding the atriopore, and anal opening (**Figures 3H–Q** and **Supplementary Video 3**). The strong endogenous GFP signal in the cirri interfered with Alexa 488 from IHC staining

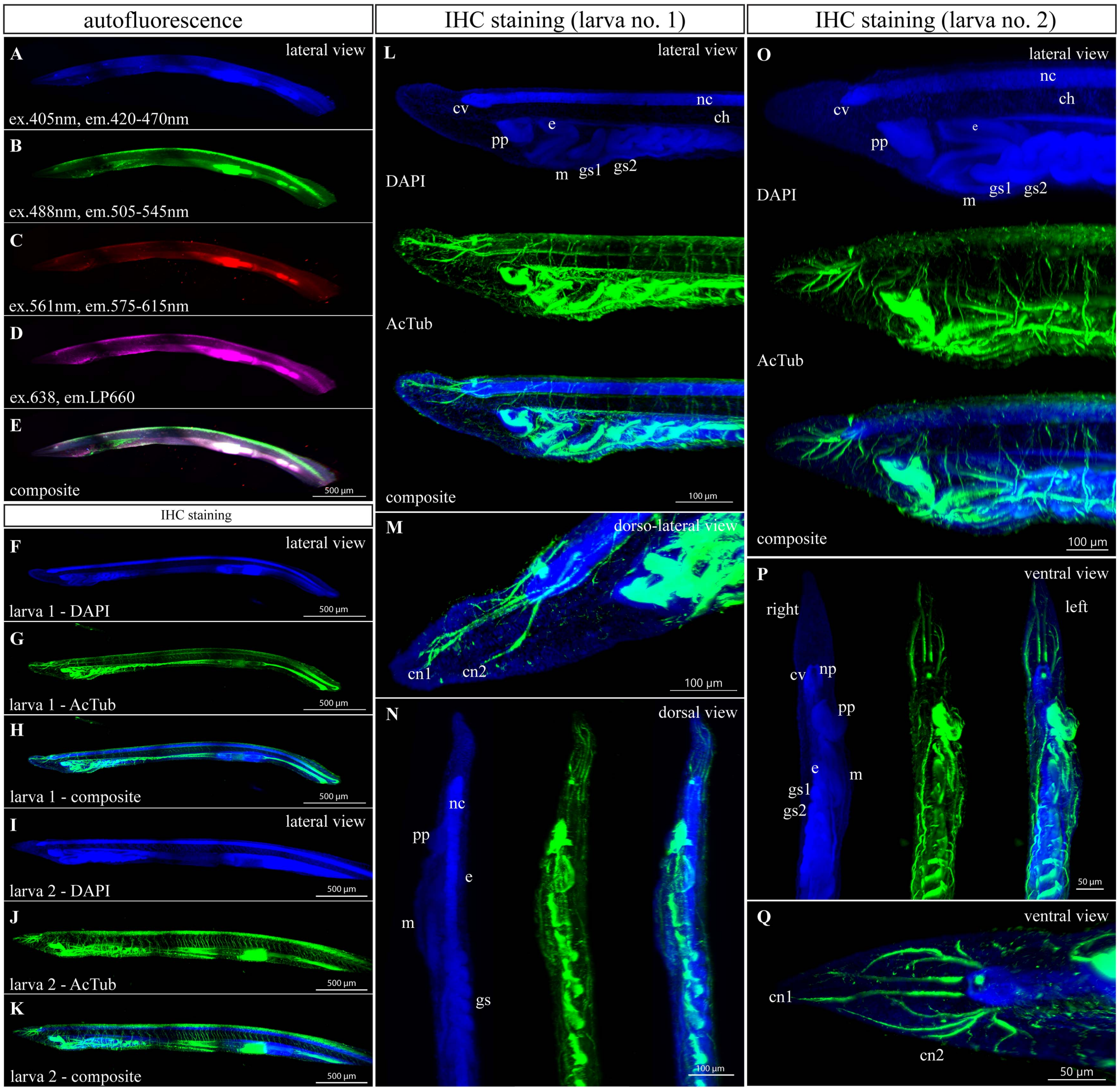

autofluorescence
IHC staining (larva no. 1)
IHC staining (larva no. 2)
A
lateral view
ex.405nm, em.420-470nm
B
ex.488nm, em.505-545nm
C
ex.561nm, em.575-615nm
D
ex.638, em.LP660
E
composite
500 µm
IHC staining
F
lateral view
larva 1 - DAPI
500 µm
G
larva 1 - AcTub
500 µm
H
larva 1 - composite
500 µm
I
lateral view
larva 2 - DAPI
500 µm
J
larva 2 - AcTub
500 µm
K
larva 2 - composite
500 µm
L
lateral view
cv
nc
ch
e
pp
m
gs1
gs2
DAPI
AcTub
composite
100 µm
M
dorso-lateral view
cn1
cn2
100 µm
N
dorsal view
nc
pp
e
m
gs
100 µm
O
lateral view
nc
ch
cv
pp
e
m
gs1
gs2
DAPI
AcTub
composite
100 µm
P
ventral view
right
left
cv
np
pp
e
m
gs1
gs2
50 µm
Q
ventral view
cn1
cn2
50 µm

FIGURE 2 | Autofluorescence and anti-acetylated-α-tubulin staining of a 1-month-old amphioxus larva. **(A–D)** Individual scan of larva autofluorescence emitted in blue **(A)**, green **(B)**, red **(C)**, and magenta channels **(D)**. **(E)** Composite image of previous scans. The larva scanned for autofluorescence was about 2.9 mm long and had seven unpaired gill slits. **(F–Q)** Immunofluorescent staining of acetylated tubulin in 1-month-old larvae. **(F–K)** Whole-mount images of two different individuals. Larva 1 was about 2.7 mm long and had seven unpaired gill slits. Larva 2 was about 3.5 mm long and had eight unpaired gill slits. **(L)** Detail of the anterior portion of larva 1 from the left side view. **(M)** Detail of cranial nerves of larva 1 from the dorsolateral view. **(N)** Detail of the anterior portion of larva 1 from the dorsal view. **(O)** Detail of the anterior portion of larva 2 from the left view. **(P)** Detail of the anterior portion of larva 2 from the ventral view. **(Q)** Detail of cranial nerves and a cerebral vesicle of larva 2 from the ventral view. Larvae 1 and 2 came from the same batch and were raised in the same conditions. All fluorescence images represent 3D reconstruction of the whole body mass. The regular areas of low signal intensity are caused by lower intensity of light falling on the camera chip margins during the tile imaging. ch., notochord; c.n.1/2, first and second cranial nerves; c.v., cerebral vesicle; e., endostyle; g.s.1/2, first and second gill slit; m., mouth; n.c., neural cord; np., neuroporus; p.p., preoral pit. Individual channels were imaged sequentially in the track switch mode—z-stack.

(**Figures 3J–M** and **Supplementary Video 3**). Therefore, when researchers are interested in studying the oral apparatus in juvenile or adult amphioxus, they should use a secondary antibody with a red or far red fluorophore.

Autofluorescence and Acetylated Tubulin Staining of a 6-Month-Old Juvenile

Next, we examined the autofluorescence of a 6-month-old specimen with 34 pairs of gill slits and length of 10.5 mm (**Figure 4**). The autofluorescence signals of the 3- and 6-month-old juveniles demonstrate many similarities. As in the case of younger juveniles, DAPI staining was performed to get the outline of the whole-mount specimen (**Figure 4A**). The green autofluorescence was the strongest in buccal cirri (**Figure 4B**). The signal emitted in the red spectrum showed weaker intensity and marked the notochord, buccal cirri, branchial apparatus, gut, and hepatic cecum (**Figure 4C**). The most obvious autofluorescence in the far-red channel was present

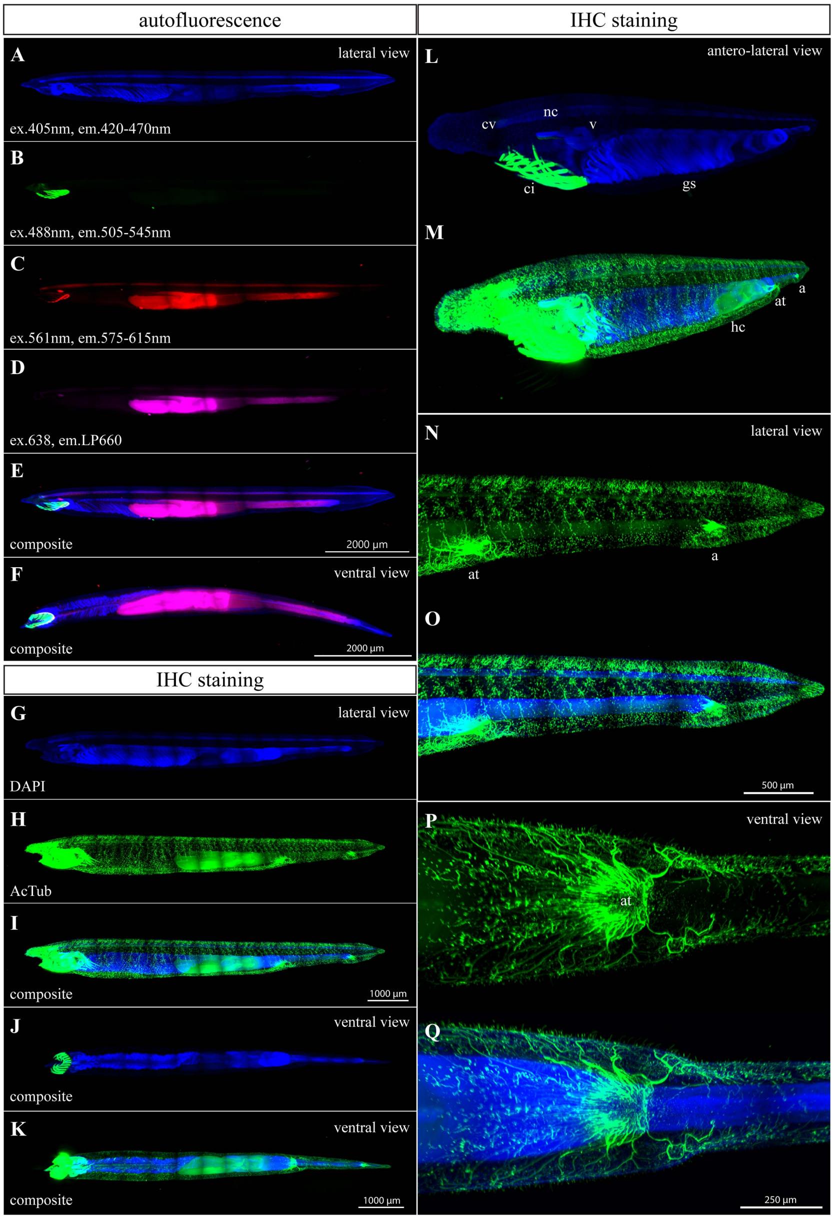

autofluorescence
IHC staining
A
lateral view
ex.405nm, em.420-470nm
B
ex.488nm, em.505-545nm
C
ex.561nm, em.575-615nm
D
ex.638, em.LP660
E
composite
2000 µm
F
ventral view
composite
2000 µm
IHC staining
G
lateral view
DAPI
H
AcTub
I
composite
1000 µm
J
ventral view
composite
K
ventral view
composite
1000 µm
L
antero-lateral view
cv
nc
v
ci
gs
M
a
at
hc
N
lateral view
at
a
O
500 µm
P
ventral view
at
Q
250 µm

FIGURE 3 | Autofluorescence and anti-acetylated-α-tubulin staining of a 3-month-old amphioxus juvenile. **(A)** DAPI staining of a specimen examined for autofluorescence. **(B–F)** Individual scans of juvenile autofluorescence emitted in green **(B)**, red **(C)**, and magenta channels **(D)**. **(E,F)** Composite image of previous scans from the lateral **(E)** and ventral view **(F)**. The specimen scanned for autofluorescence was about 8 mm long. **(G)** DAPI staining of the specimen in **(H–I)**. **(H–N)** IHC staining of acetylated tubulin. The stained juvenile was 8 mm long. **(J,K)** Ventral view of immunostained juvenile. **(L,M)** Anterolateral view of the anterior portion of immunostained juvenile. Differential adjustment of brightness/contrast of the green channel was required to demonstrate the intensity of strong endogenous GFP signal in cirri **(L)** compared with the staining of acetylated tubulin **(M)**. **(N,O)** Detail of the posterior portion of immunostained juvenile, from the lateral view. **(P,Q)** Ventral view of atriopore innervation. All fluorescence images represent 3D reconstruction of the whole body mass. The regular areas of low signal intensity in the images are caused by lower intensity of the light falling on the camera chip margins during the imaging of the tiles. a., anus; at., atriopore; ci., buccal cirri; c.v., cerebral vesicle; g.s., gill slits; h.c., hepatic cecum; n.c., neural cord; v., velum.

in the hepatic cecum and gut (**Figure 4D**). Remarkably, optical sectioning revealed a strong signal in the notochord that could be used as endogenous marker (**Figures 4F,G**).

The immunostained specimen with the length of 9.5 mm came from the same batch. The immunostaining of acetylated tubulin visualized the increasing complexity of neural organization of the growing amphioxus juvenile (**Figures 4H–Q** and **Supplementary Videos 4, 5**). Imaging of cleared specimens allowed us to observe distinct cranial or spinal nerves rising from the neural cord, including branches of the second and third spinal nerves that innervate the right and left side of the velum (**Figures 4M,N,P,Q**). Interestingly, the innervation around the atriopore becomes

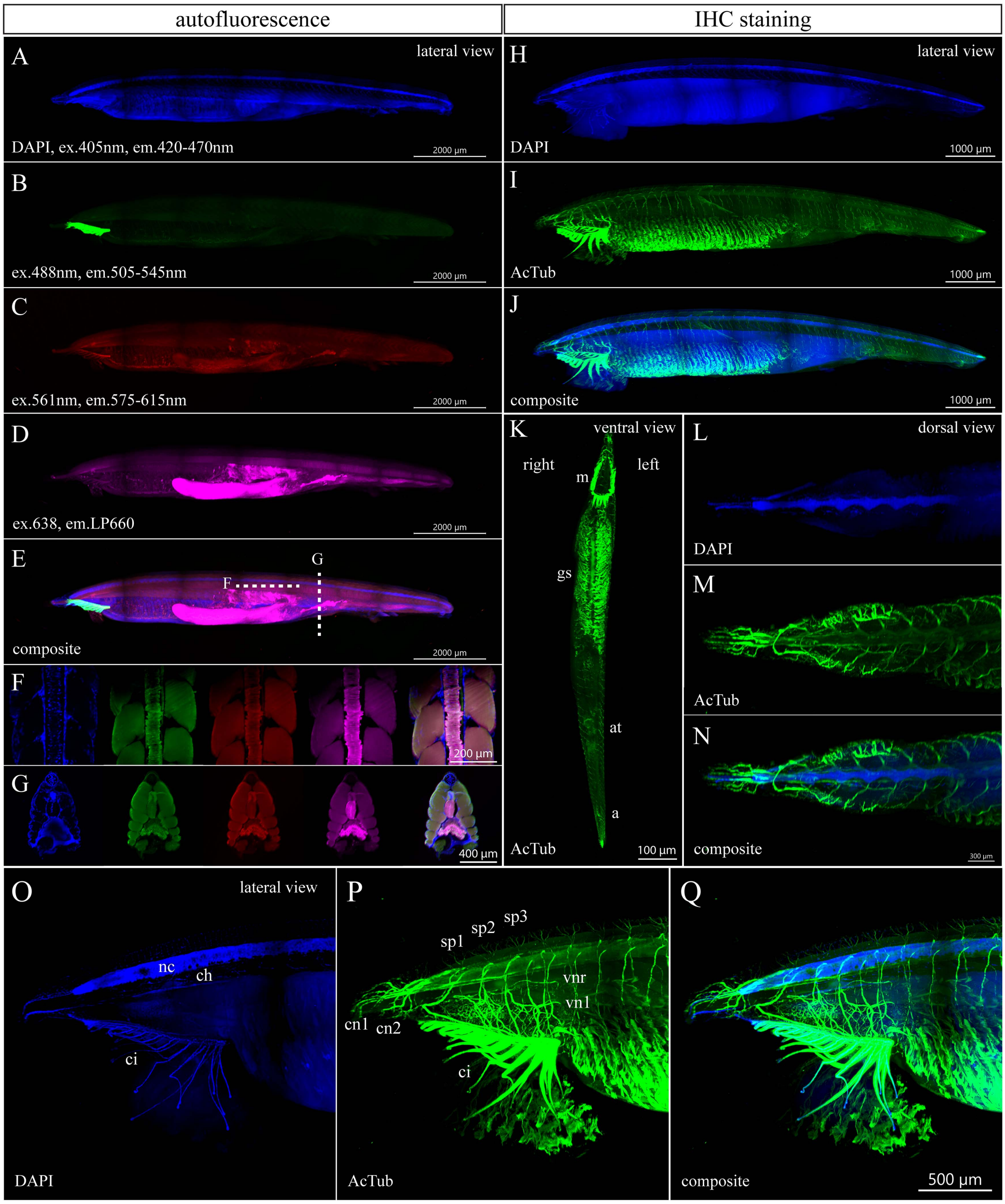

autofluorescence
IHC staining
A
lateral view
DAPI, ex.405nm, em.420-470nm
2000 μm
B
ex.488nm, em.505-545nm
2000 μm
C
ex.561nm, em.575-615nm
2000 μm
D
ex.638, em.LP660
2000 μm
E
F
G
composite
2000 μm
F
200 μm
G
400 μm
H
lateral view
DAPI
1000 μm
I
AcTub
1000 μm
J
composite
1000 μm
K
ventral view
right
left
m
gs
at
a
AcTub
100 μm
L
dorsal view
DAPI
M
AcTub
N
composite
300 μm
O
lateral view
nc
ch
ci
DAPI
P
sp1
sp2
sp3
vnr
vnl
cn1
cn2
ci
AcTub
Q
composite
500 μm

FIGURE 4 | Autofluorescence and anti-acetylated-α-tubulin staining of a 6-month-old amphioxus juvenile. **(A)** DAPI staining of a 6-month-old juvenile with a length of 10.5 mm. **(B–E)** Autofluorescence of the same individual emitted in green **(B)**, red **(C)**, and magenta channels **(D)**. **(E)** Composite image from the lateral view. **(F,G)** Subset of horizontal **(F)** or transversal **(G)** optical slices of the specimen from **(A–E)** in the area with notochord showing the DAPI staining and autofluorescence in green, red, and magenta channels with a composite of all above images (from left to right). **(H–Q)** IHC staining of acetylated tubulin of a 6-month-old juvenile with a length of 9.7 mm shown as whole mount **(H–K)** or individual portion of the body **(L–Q)**. Detail of the anterior portion of juvenile body after removal of optical slices containing buccal cirri, **(O–Q)** detail of the anterior juvenile body part from the left side view. The regular areas of low signal intensity in the images are caused by lower light intensity falling on the camera chip margins during the imaging of the tiles. a., anus; at., atrioporus; ci., buccal cirri; ch., notochord; c.n.1/2, first and second cranial nerve; g.s., gill slits; m., mouth; n.c., neural cord; sp.1/2/3, first, second, and third spinal nerves; v.n.l./v.n.r., nerve to the left/right side of the velum. Individual channels were imaged sequentially in the track switch mode—z-stack.

profoundly reduced in the 6-month-old juvenile as compared with the 3-month-old specimen (compare **Figures 3**, **4**). Taken together, the whole-mount IHC of cleared amphioxus larvae and juvenile specimens with anti-acetylated-α-tubulin antibodies allowed us to reveal the details of the developing amphioxus nerve net.

Immunohistochemistry Staining of an Amphioxus 1-Month-Old Larva and a 3-Month-Old Juvenile With Anti-Melanopsin Antibody

Next, we tested the possibility to perform whole-mount immunostaining of larger amphioxus specimens with antibodies generated against amphioxus melanopsin. As previously described, melanopsin is present in rhabdomeric photoreceptive cells (Joseph and Hesse cells) in amphioxus adults (Koyanagi et al., 2005; Pergner and Kozmik, 2017). We were able to detect the signal in Hesse photoreceptive cells in a 1-month-old larva and a 3-month-old juvenile (**Figure 5**). In contrast to the acetylated-α-tubulin, the intensity of melanopsin staining was weaker and required higher power of excitation laser. In addition, an undesired ectopic signal caused by non-specific sticking of the antibodies was present on the whole surface of the specimens (**Figures 5A–C,F**). This phenomenon was previously observed in the case of larva staining with many other antibodies (Vopalensky et al., 2012; Pergner et al., 2020). The advantage of the imaging technique based on optical sectioning lies in the possibility to observe the cells or structures hidden in the context of 3D whole mount. We visualized Hesse cells in the neural tube by removing the optical sections with skin tissue (**Figures 5C–H**). We were unable to detect melanopsin in the 6-month-old juveniles, most likely due to the low permeabilization of the dense tissue.

DISCUSSION

In this study, we demonstrated that light sheet microscopy (LSM) combined with tissue clearing enables rapid imaging of large amphioxus specimens with single-cell resolution. The main advantage of the whole-mount scanning is the possibility to examine large specimens while the tissue context is preserved. In addition, compared with confocal fluorescence microscopy, LSM imaging is faster and avoids photobleaching of deeper layers during the scanning of thick samples (Keller and Stelzer, 2008; Kromm et al., 2016).

Staining of acetylated tubulin in the developing amphioxus revealed detailed nerve branching that could be observed in the context of the whole mount and as individual sections at high resolution. Anti-acetylated tubulin antibody is frequently used by researchers to visualize the nerve net in vertebrates and invertebrates. In amphioxus, the staining was previously applied to describe the peripheral nervous system of early larvae, premetamorphic, metamorphosing, and advanced postmetamorphosis individuals (Kaji et al., 2001; Yasui and Kaji, 2008; Kaji et al., 2009; Yasui et al., 2014; Kaji et al., 2016) or neural organization of the brain in the adult (Castro et al., 2015). These earlier studies performed confocal imaging of selected segments or individual paraffin sections. The cytoarchitecture of the central nervous system of late premetamorphic larva, postmetamorphic juvenile, and adult amphioxus was studied with classical histological methods and by immunohistochemistry using acetylated tubulin (Lacalli et al., 1994; Lacalli, 2000, 2003; Ekhart et al., 2003; Castro et al., 2015). In 3– and 6-month-old juveniles, we identified several cell types that have been described previously, including anterolateral migrated cells, transluminal cells, somatic motoneurons, lamellate cells, and Retzius bipolar cells (**Supplementary Figure 1**).

Compared with acetylated tubulin, the whole-mount staining of melanopsin was weaker, and we were unable to detect any signal in 6-month-old juveniles. Generally, detection of the signal in internal structures within whole-mount staining of older specimens is always challenging due to the high density of the sample. Prolonging the incubation time with antibody together with permeabilization methods may help to overcome this obstacle. In amphioxus, melanopsin is expressed in photoreceptor cells of the Hesse organ (also called dorsal ocelli) and Joseph cells (Koyanagi et al., 2005; Pergner and Kozmik, 2017). While the first dorsal ocelli are already developing in the ventral part of the neural tube at the mid-neurula stage, Joseph cells have so far been identified in the dorso-caudal region of the cerebral vesicle only in the adult amphioxus (Ruiz and Anadon, 1991; Lacalli et al., 1994; Koyanagi et al., 2005; Castro et al., 2015). In 3-month-old juveniles, we detected melanopsin only in the dorsal ocelli, which is consistent with the idea that Joseph cells develop later during amphioxus ontogenesis.

Autofluorescence often complicates the analysis of fluorescently stained specimens. The amphioxus genome encodes a variety of GFP-like proteins with distinct spectral properties (Deheyn et al., 2007; Bomati et al., 2009; Yue et al., 2016). Although the CUBIC clearing procedure preserves the signal of endogenous fluorophores, the intensity of autofluorescence may be affected by any chemicals during fixation and staining. In this study, we observed a weak signal of endogenous fluorescence across the spectrum with the strongest one in the green channel. The high variability in the autofluorescence patterns among the different developmental stages is intriguing. Bomati et al. (2009) identified 16 GFP-like genes in the amphioxus genome belonging to six clades. Individual GFPs not only possess different spectral characteristics but their expression is also highly dynamic during development. Besides, it was shown that GFPa1 is selectively expressed in the oral cirri of the adults suggesting that some GFPs may also display the tissue-specific expression (Deheyn et al., 2007). The absorbance spectrum among different GFPs ranges from 375 to about 525 nm with a peak between 470 and 504 nm. The weak autofluorescence in the blue spectrum may be explained by the ability of some GFPs to be excited at 405 nm. Detailed description of the scanning parameters and characterization of the autofluorescence of distinct structures may help researchers to choose the secondary antibody with a suitable fluorophore. At the same time, endogenous fluorescence can be used as counterstaining or as a marker to observe the cells or structures of a specimen without additional staining. In fact, the autofluorescence of individual structures may be used not only to study the intact animals but also to analyze the phenotypes after gene perturbations or chemical treatment.

For example, the notochord of 3– and 6-month-old juveniles is nicely visualized by setting the excitation wavelength at 638 nm. The parameters of notochord autofluorescence correlate with the spectral properties of porphyrins. Porphyrins are present

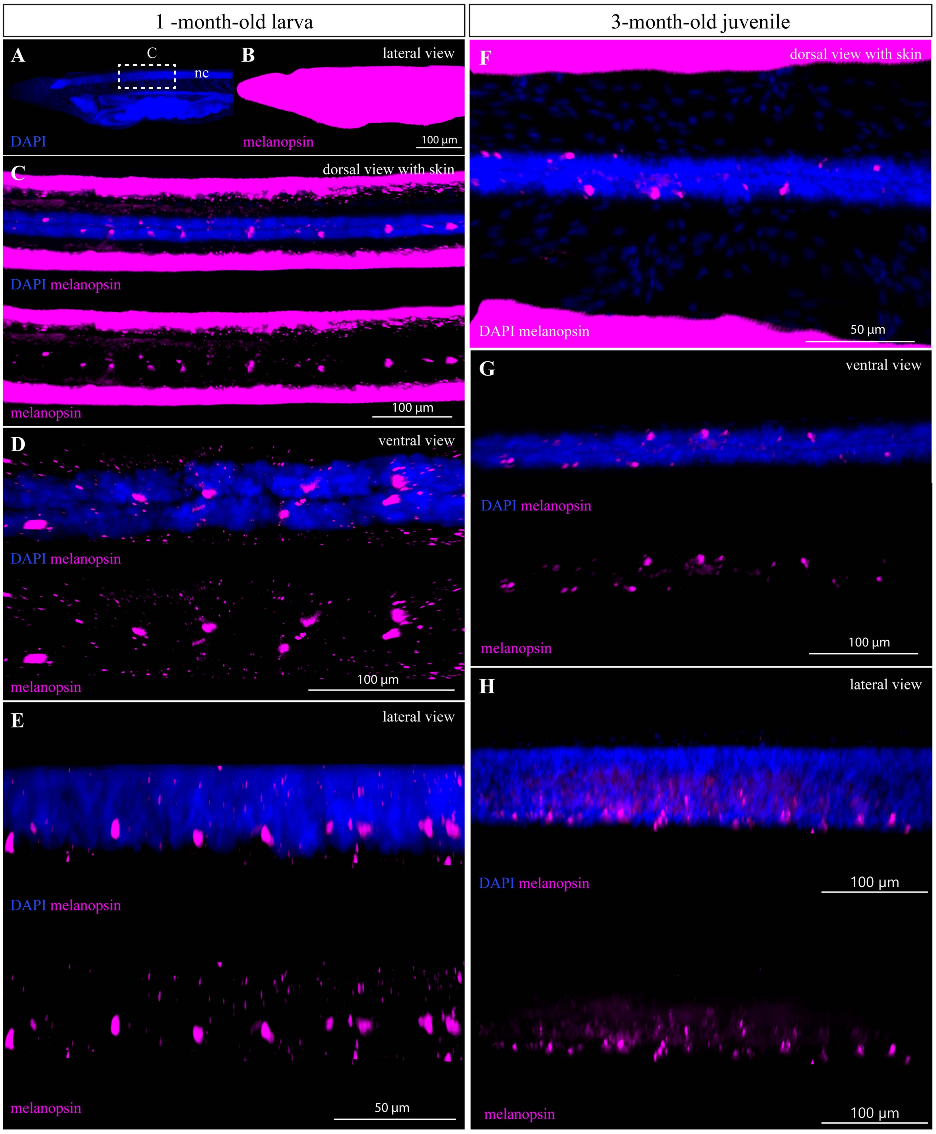

1 -month-old larva
3-month-old juvenile
A
C
nc
DAPI
B
lateral view
melanopsin
100 µm
C
dorsal view with skin
DAPI melanopsin
melanopsin
100 µm
D
ventral view
DAPI melanopsin
melanopsin
100 µm
E
lateral view
DAPI melanopsin
melanopsin
50 µm
F
dorsal view with skin
DAPI melanopsin
50 µm
G
ventral view
DAPI melanopsin
melanopsin
100 µm
H
lateral view
DAPI melanopsin
100 µm
melanopsin
100 µm

FIGURE 5 | IHC staining of melanopsin in 1- and 3-month-old amphioxus specimens. **(A,B)** Whole mount of the anterior part of 1-month-old larva shown with the adjustment of brightness/contrast required to visualize the signal of melanopsin in the neural tube **(C–E)**. Dotted rectangle marks the region from which the neural tube was cropped by processing in IMARIS analysis software shown in **(C–E)**. **(C)** Subset of horizontal slices containing ventral neural cord with melanopsin-positive cells. **(D)** Subset of horizontal slices containing neural cord and melanopsin-positive cells after removal of optical slices containing skin tissue by the processing software. **(E)** Cropped area containing the neural cord with melanopsin-positive cells (with and without DAPI staining) from the lateral view. **(F)** Subset of horizontal slices with the neural tube of a 3-month-old juvenile stained with DAPI and melanopsin antibody. **(G)** Subset of horizontal slices containing the neural cord and melanopsin-positive cells after optical removal of skin tissue, from the dorsal view. **(H)** Detail of neural cord with melanopsin-positive cells (with and without DAPI staining) after removal of optical slices containing overstained skin tissue, from the lateral view. The specimens were imaged sequentially in the track switch mode—z-stack.

in different tissues and exhibit a broad absorption spectra with one peak around 400 nm and four peaks between 490 and 650 nm. The emission spectra of porphyrins are located between 550 and 700 nm (Plus, 1992). Porphyrins constitute a part of hemoglobin protein subunits. The presence of hemoglobin in the notochord of amphioxus was demonstrated by high-performance liquid chromatography and by measuring optical absorption spectra (Bishop et al., 1998). The range of absorption spectra of the notochord sample was from 400 to 650 nm. We therefore speculate that the observed notochord autofluorescence may be caused by the presence of hemoglobin or some other proteins associated with porphyrins.

An important issue in SPIM imaging is the production of large amounts of data and, thus, the need for high-performance computational technology and large data storage capabilities. In addition, higher amounts of data require more time for their processing. Therefore, careful planning of the experiment in terms of the used resolution, zoom, volume, and number of channels should be performed before the scanning itself (Amat et al., 2015; Kromm et al., 2016). For example, the whole mount of the specimen can be imaged at a lower zoom level and the details with higher magnification (**Supplementary Table 1**).

Taken together, in this study, we demonstrated that LSM may be applied to study the whole mount of large amphioxus specimens at high resolution. We believe that this technique will constitute a valuable complement to the conventional methods of electron (Lacalli et al., 1994, 1999; Wicht and Lacalli, 2005) and confocal microscopy (Yasui and Kaji, 2008). Moreover, the possibility to visualize the entire amphioxus body at the single-cell resolution may facilitate unbiased identification of the morphological and molecular differences between wild-type and experimentally manipulated animals.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SM, HC, ZK, and IK designed the study and conceived the experiments. SM, HC, and AV performed the wet lab experiments. SM, HC, AV, and IK analyzed the data. IK and ZK provided new reagents and animal specimens. SM and IK wrote the manuscript. All authors have read and approved the manuscript.

FUNDING

This work was supported by the Czech Science Foundation grant 20-25377S awarded to ZK.

ACKNOWLEDGMENTS

We are grateful to Veronika Noskova for the amphioxus facility maintenance. We acknowledge the Light Microscopy Core Facility, IMG CAS, Prague, Czech Republic, supported by MEYS (LM2018129, CZ.02.1.01/0.0/0.0/18_046/0016045) and RVO: 68378050-KAV-NPUI, for their support with the light sheet microscopy presented herein.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.702986/full#supplementary-material

Supplementary Figure 1 | Selected optical sections show different types of neurons immunostained with anti-acetylated tubulin in 3-month-old juvenile **(A-C")** and 6-month-old juvenile **(D-E")**. **(A-A")** Anteriolateral migrated cells. **(B-B")** Translumenal cells. **(C-C")** Somatic motoneurons. **(D-D")** Lamellate cells. **(E-E")** Retzius bipolar cells. Arrows in the individual panel point to indicated cell type. Faint vertical and horizontal lines indicate the plane of section.

Supplementary Table 1 | Technical parameters of the imaging of amphioxus specimens.

Supplementary Table 2 | Technical specification of analysis computers.

Supplementary Video 1 | X-/y-axis rotation of 3D reconstruction of 1-month-old larva immunostained with anti-acetylated tubulin (green) antibody with DAPI (blue).

Supplementary Video 2 | Removal of optical horizontal slices out of 1-month-old larva immunostained with anti-acetylated tubulin (green) antibody with DAPI (blue).

Supplementary Video 3 | Rotation of 3D whole mount and close-ups of the parts of 3-month-old larva immunostained with anti-acetylated tubulin (green) antibody with DAPI (blue). (0:00:00–0:00:10 s) X-/y-axis rotation of the 3D whole mount. (0:00:11–0:00:20 s) buccal cirri highlighted by the adjustment of the brightness/contrast. (0:00:20–0:00:50 s) slow-motion focusing on details of the nerve net from the lateral view.

Supplementary Video 4 | Immunostaining of 6-month-old juvenile I with anti-acetylated tubulin antibody (green) and DAPI (blue). (0:00:00–0:00:36 s) 3D whole mount and slow-motion of the parts of the specimen from the lateral view. (0:00:37–0:00 s) Slow-motion of 3D whole mount from the ventral view.

Supplementary Video 5 | Immunostaining of 6month-old juvenile II with anti-acetylated tubulin antibody (green) and DAPI (blue). (0:00:00–0:00:05 s) Rotation of 3D whole mount to the ventral view (0:00:06–0:00:18 s) Slow-motion of 3D whole mount from the ventral view (0:00:19–0:00:36 s) Slow-motion of the specimen from the dorsal view gradually rotated to the lateral view after removal of ventral slices by image processing. (0:00:37–0:00:50 s) X-/z-rotation of 3D whole mount from the lateral to frontal view.

REFERENCES

Amat, F., Höckendorf, B., Wan, Y., Lemon, W. C., McDole, K., and Keller, P. J. (2015). Efficient processing and analysis of large-scale light-sheet microscopy data. *Nat. Protoc.* 10, 1679–1696. doi: 10.1038/nprot.2015.111

Bishop, J. J., Vandergon, T. L., Green, D. B., Doeller, J. E., and Kraus, D. W. (1998). A High-Affinity Hemoglobin Is Expressed in the Notochord of Amphioxus, Branchiostoma californiense. *Biol. Bull.* 195, 255–259. doi: 10.2307/1543136

Bomati, E. K., Manning, G., and Deheyn, D. D. (2009). Amphioxus encodes the largest known family of green fluorescent proteins, which have diversified into distinct functional classes. *BMC Evol. Biol.* 9:77. doi: 10.1186/1471-2148-9-77

Bozzo, M., Pergner, J., Kozmik, Z., and Kozmikova, I. (2017). Novel polyclonal antibodies as a useful tool for expression studies in amphioxus embryos. *Int. J. Dev. Biol.* 61, 793–800. doi: 10.1387/ijdb.170259ik

Castro, A., Becerra, M., Manso, M. J., and Anadón, R. (2015). Neuronal organization of the brain in the adult amphioxus (Branchiostoma lanceolatum): a study with acetylated tubulin immunohistochemistry. *J. Comp. Neurol.* 523, 2211–2232. doi: 10.1002/cne.23785

Conklin, E. G. (1932). The embryology of amphioxus. *J. Morphol.* 54, 69–151. doi: 10.1002/jmor.1050540103

Deheyn, D. D., Kubokawa, K., Mccarthy, J. K., Murakami, A., Porrachia, M., Rouse, G. W., et al. (2007). Endogenous green fluorescent protein (GFP) in amphioxus. *Biol. Bull.* 213, 95–100. doi: 10.2307/25066625

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. doi: 10.1038/nature04336

Ekhart, D., Korf, H. W., and Wicht, H. (2003). Cytoarchitecture, topography, and descending supraspinal projections in the anterior central nervous system of Branchiostoma lanceolatum. *J. Comp. Neurol.* 466, 319–330. doi: 10.1002/cne.10803

Fu, Q., Martin, B. L., Matus, D. Q., and Gao, L. (2016). Imaging multicellular specimens with real-time optimized tiling light-sheet selective plane illumination microscopy. *Nat. Commun.* 7:11088. doi: 10.1038/ncomms11088

Hatschek, B. (1881). *Studien über Entwicklung des Amphioxus*. Ann Arbor: University of Michigan Library.

Huisken, J., and Stainier, D. Y. R. (2009). Selective plane illumination microscopy techniques in developmental biology. *Development* 136, 1963–1975. doi: 10.1242/dev.022426

Kaji, T., Aizawa, S., Uemura, M., and Yasui, K. (2001). Establishment of left-right asymmetric innervation in the lancelet oral region. *J. Comp. Neurol.* 435, 394–405. doi: 10.1002/cne.1039

Kaji, T., Reimer, J. D., Morov, A. R., Kuratani, S., and Yasui, K. (2016). Amphioxus mouth after dorso-ventral inversion. *Zool. Lett.* 2:2. doi: 10.1186/s40851-016-0038-3

Kaji, T., Shimizu, K., Artinger, K. B., and Yasui, K. (2009). Dynamic modification of oral innervation during metamorphosis in branchiostoma belcheri, the oriental lancelet. *Biol. Bull.* 217, 151–160. doi: 10.1086/BBLv217n2p151

Keller, P. J., and Stelzer, E. H. (2008). Quantitative in vivo imaging of entire embryos with Digital Scanned Laser Light Sheet Fluorescence Microscopy. *Curr. Opin. Neurobiol.* 18, 624–632. doi: 10.1016/j.conb.2009.03.008

Koyanagi, M., Kubokawa, K., Tsukamoto, H., Shichida, Y., and Terakita, A. (2005). Cephalochordate melanopsin: evolutionary linkage between invertebrate visual cells and vertebrate photosensitive retinal ganglion cells. *Curr. Biol.* 15, 1065–1069. doi: 10.1016/j.cub.2005.04.063

Kromm, D., Thumberger, T., and Wittbrodt, J. (2016). *An Eye On Light-Sheet Microscopy*. Netherlands: Elsevier Ltd, doi: 10.1016/bs.mcb.2016.01.001

Lacalli, T. C. (2000). Cell morphology in amphioxus nerve cord may reflect the time course of cell differentiation. *Int. J. Dev. Biol.* 44, 903–906.

Lacalli, T. C. (2003). Ventral neurons in the anterior nerve cord of amphioxus larvae. II. Further data on the pacemaker circuit. *J. Morphol.* 257, 212–218. doi: 10.1002/jmor.10133

Lacalli, T. C., Gilmour, T. H. J., and Kelly, S. J. (1999). The oral nerve plexus in amphioxus larvae: function, cell types and phylogenetic significance. *Proc. R. Soc. B Biol. Sci.* 266, 1461–1470. doi: 10.1098/rspb.1999.0801

Lacalli, T. C., Holland, N. D., and West, J. E. (1994). Landmarks in the anterior central nervous system of amphioxus larvae. *Philos. Trans. R. Soc. B Biol. Sci.* 344, 165–185. doi: 10.1098/rstb.1994.0059

Masson, A., Escande, P., Frongia, C., Clouvel, G., Ducommun, B., and Lorenzo, C. (2015). High-resolution in-depth imaging of optically cleared thick samples using an adaptive SPIM. *Sci. Rep.* 5:16898. doi: 10.1038/srep16898

Nie, J., Liu, S., Yu, T., Li, Y., Ping, J., Wan, P., et al. (2020). Fast, 3D Isotropic Imaging of Whole Mouse Brain Using Multiangle-Resolved Subvoxel SPIM. *Adv. Sci.* 7:1901891. doi: 10.1002/advs.201901891

Pergner, J., and Kozmik, Z. (2017). Amphioxus photoreceptors - Insights into the evolution of vertebrate opsins, vision and circadian rhythmicity. *Int. J. Dev. Biol.* 61, 665–681. doi: 10.1387/ijdb.170230zk

Pergner, J., Vavrova, A., Kozmikova, I., and Kozmik, Z. (2020). Molecular Fingerprint of Amphioxus Frontal Eye Illuminates the Evolution of Homologous Cell Types in the Chordate Retina. *Front. Cell Dev. Biol.* 8:705. doi: 10.3389/fcell.2020.00705

Plus, R. (1992). A review of in vivo studies of porphyrins and unexpected fluorescences. An interpretation of the results. *Med. Hypotheses* 37, 49–57. doi: 10.1016/0306-9877(92)90014-4

Ruiz, M. S., and Anadon, R. (1991). Some considerations on the fine structure of rhabdomeric photoreceptors in the amphioxus, Branchiostoma lanceolatum (*cephalochordata*). *J. Hirnforsch.* 32, 159–164.

Stokes, M. D., and Holland, N. D. (1995). Embryos and Larvae of a Lancelet, Branchiostoma floridae, from Hatching through Metamorphosis: growth in the Laboratory and External Morphology. *Acta Zool.* 76, 105–120. doi: 10.1111/j.1463-6395.1995.tb00986.x

Susaki, E. A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A., and Ueda, H. R. (2015). Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat. Protoc.* 10, 1709–1727. doi: 10.1038/nprot.2015.085

Vopalensky, P., Pergner, J., Liegertova, M., Benito-Gutierrez, E., Arendt, D., and Kozmik, Z. (2012). Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. *Proc. Natl. Acad. Sci. U. S. A.* 109, 15383–15388. doi: 10.1073/pnas.1207580109

Wicht, H., and Lacalli, T. C. (2005). The nervous system of amphioxus: structure, development, and evolutionary significance. *Can. J. Zool.* 83, 122–150. doi: 10.1139/z04-163

Willey, A. (1891). The later larval development of amphioxus. *J. Cell Sci.* 32, 183–234. doi: 10.1038/044202a0

Willey, A. (1894). *Amphioxus And The Ancestry Of The Vertebrates*. New York and London: Macmillan and Co, doi: 10.1038/051433a0

Yasui, K., and Kaji, T. (2008). The lancelet and ammocoete mouths. *Zoolog. Sci.* 25, 1012–1019. doi: 10.2108/zsj.25.1012

Yasui, K., Kaji, T., Morov, A. R., and Yonemura, S. (2014). Development of oral and branchial muscles in lancelet larvae of Branchiostoma japonicum. *J. Morphol.* 275, 465–477. doi: 10.1002/jmor.20228

Yue, J. X., Holland, N. D., Holland, L. Z., and Deheyn, D. D. (2016). The evolution of genes encoding for green fluorescent proteins: insights from cephalochordates (amphioxus). *Sci. Rep.* 6:28350. doi: 10.1038/srep28350

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Machacova, Chmelova, Vavrova, Kozmik and Kozmikova. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 03 August 2021
doi: 10.3389/fcell.2021.709696

The Ascidia *Ciona robusta* Provides Novel Insights on the Evolution of the AP-1 Transcriptional Complex

*Pina Marotta[1,2], Federica Salatiello[3], Luca Ambrosino[2], Federica Berruto[3], Maria Luisa Chiusano[2,4] and Annamaria Locascio[3]**

[1] Stazione Zoologica Anton Dohrn, Department of Integrative Marine Ecology, Naples, Italy, [2] Stazione Zoologica Anton Dohrn, Department of Research Infrastructures for Marine Biological Resources, Naples, Italy, [3] Stazione Zoologica Anton Dohrn, Department of Biology and Evolution of Marine Organisms, Naples, Italy, [4] Department of Agriculture, Università degli Studi di Napoli Federico II, Portici, Italy

The Activator Protein-1 transcription factor family (AP-1) transcriptional complex is historically defined as an early response group of transcription factors formed by dimeric complexes of the Jun, Fos, Atf, and Maf bZIP proteins that control cell proliferation and differentiation by regulating gene expression. It has been greatly investigated in many model organisms across metazoan evolution. Nevertheless, its complexity and variability of action made its multiple functions difficult to be defined. Here, we place the foundations for understanding the complexity of AP-1 transcriptional members in tunicates. We investigated the gene members of this family in the ascidian *Ciona robusta* and identified single copies of Jun, Fos, Atf3, Atf2/7, and Maf bZIP-related factors that could have a role in the formation of the AP-1 complex. We highlight that mesenchyme is a common cellular population where all these factors are expressed during embryonic development, and that, moreover, *Fos* shows a wider pattern of expression including also notochord and neural cells. By ectopic expression in transgenic embryos of *Jun* and *Fos* genes alone or in combination, we investigated the phenotypic alterations induced by these factors and highlighted a degree of functional conservation of the AP-1 complex between *Ciona* and vertebrates. The lack of gene redundancy and the first pieces of evidence of conserved functions in the control of cell movements and structural organization exerted by these factors open the way for using *Ciona* as a helpful model system to uncover the multiple potentialities of this highly complex family of bZIP transcription factors.

Keywords: Jun, Fos, bZIP protein, mesenchyme, transcription factor, notochord

OPEN ACCESS

Edited by:
Stephanie Bertrand,
UMR 7232 Biologie Intégrative des Organismes Marins (BIOM), France

Reviewed by:
Pedro Martinez,
University of Barcelona, Spain
Simona Candiani,
University of Genoa, Italy

****Correspondence:***
Annamaria Locascio
annamaria.locascio@szn.it

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *14 May 2021*
Accepted: *12 July 2021*
Published: *03 August 2021*

Citation:
Marotta P, Salatiello F, Ambrosino L, Berruto F, Chiusano ML and Locascio A (2021) The Ascidia Ciona robusta Provides Novel Insights on the Evolution of the AP-1 Transcriptional Complex. Front. Cell Dev. Biol. 9:709696. doi: 10.3389/fcell.2021.709696

INTRODUCTION

It is now commonly recognized that the complexity of organisms does not always directly correlate with the number of genes in its genome. Genetic diversity is often ensured by more sophisticated control and interactions among the already available factors (Levine and Tjian, 2003; Van Nimwegen, 2003; Ranea et al., 2005). One of the best ways of creating a wider possibility of regulatory responses is through the formation of transcription factor (TF) dimers, with specific DNA-binding properties, that allows the definition of novel control circuits (Klemm et al., 1998).

To understand animal evolution, it is important to delineate how DNA binding and protein interactions of dimerizing TFs evolved. Jun and Fos, belonging to the bZIP proteins, are dimerizing factors found in all eukaryotes and are the most common components of the Activator Protein-1 TF family (AP-1).

The AP-1 complex collectively describes a group of structurally and functionally related TFs of the Jun (Jun, JunB, and JunD) and Fos protein families (Fos, FosB, Fra-1, and Fra-2) and some members of the Atf (Atf, Atf-2, and Atf-3) and Jdp (Jdp-1 and Jdp-2) subfamilies. These factors share structural similarities and form homo- and heterodimeric complexes among them. Each of these proteins is differentially expressed and regulated, determining that every cell type has a complex different mixture of AP-1 dimers with different functions (reviewed by Wagner, 2001). The Jun proteins exist both as homo- and heterodimers, while the Fos proteins form stable heterodimers with Jun and the other AP-1 proteins. All the combinations of AP-1 complexes bind a consensus sequence along the DNA (5′-TGAG/CTCA3′), known as the 12-*O*–tetradecanoylphorbol-13-acetate (TPA)-responsive element (TRE) (Angel and Karin, 1991), and control both basal and inducible transcription of several genes containing the AP-1-binding sites in their regulatory sequences.

During evolution, various layers of AP-1 complexity have appeared such as gene duplications and differences between bZIP members and changes in their co-localization and interaction. This diversification, together with the emergence of different downstream target genes, greatly contributed to extend their mechanisms of action.

About AP-1 functional evolution, in *Drosophila*, Fos is crucial for dorsal closure (Zeitlinger et al., 1997), and in *Caenorhabditis elegans*, it is necessary for cell invasion through the basement membrane during vulvar development (Sherwood et al., 2005). Among vertebrates, specific roles have been assigned to c-Fos and c-Jun in skeletal and bone morphogenesis, where loss of these genes in mice causes osteopenia and defective bone remodeling (Wang et al., 1992; Kenner et al., 2004). In particular, c-Jun is required for axial skeletogenesis by regulating notochord survival and intervertebral disc formation (Behrens et al., 2003), and c-Fos is expressed in both nucleus pulposus cells (Lee et al., 2007) and chordomas (Schwab et al., 2009). Furthermore, several *in vitro* and *in vivo* studies pointed to an important role of the AP-1 complex for hematopoiesis specification in different vertebrate species, such as the mouse (Eferl et al., 1999; Lichtinger et al., 2012; Pereira et al., 2013; Goode et al., 2016; Obier et al., 2016), *Xenopus* (Lee et al., 2012), and zebrafish (Wei et al., 2014; Li et al., 2015). However, none of these studies has completely identified the whole molecular mechanisms responsible for these effects, also because of the number of AP-1 family members and their highly variable dimeric composition, that are also depending on specific cellular contexts.

To date, genetically modified mice harboring genetic disruption and/or transgenic overexpression as well as mutant cells isolated from these animals have represented the major source to understand the regulatory functions of AP-1 subunits. They have highlighted how AP-1 mediates gene regulation in response to a plethora of physiological stimuli, including cytokines, growth factors, and stress signals, in a variety of cellular events involved in normal and neoplastic development, such as proliferation, differentiation, apoptosis, and transformation (Karin et al., 1997). Moreover, as a general rule derived from all studies, the AP-1 family members must be present in a well-defined asset to coordinate and ensure proper development or physiology of the organism.

In light of the great complexity of this family of factors, investigating their function in specific cell types and animal models suitable for transgenesis and reverse genetic approaches could be very useful to broaden the knowledge on their multiple functions and mechanisms of action. *Ciona robusta* is an excellent model for genetic and embryonic studies—thanks to its rapid development, traceable cell lineage, and the availability of an efficient electroporation method for transgenic experiments. Furthermore, its phylogenetic position at the base of vertebrate origin before the 2R duplication and its complete and annotated genomic sequence pose this simple chordate to the forefront of the model organisms available for functional and regulatory studies. Most notably is that *Ciona* possesses the basic developmental features of vertebrates driven by non-duplicated gene families.

The ascidian bZIP factors have been identified and classified by Yamada et al. (2003) in 2003. In this work, all the known human bZIP genes were confirmed to have similarity vs. both the *C. robusta* genome and its cDNA/EST collections. With this analysis, 26 candidate genes were identified representing an almost complete set of bZIP genes in *Ciona*, and among them, one *Jun* gene and one *Fos* gene were identified, respectively (Yamada et al., 2003). Next, Imai et al. (2004) analyzed the expression profile of most of the ascidian TFs during *Ciona* embryonic development. They found that *Jun* and *Fos* are expressed in the B7.7 mesenchyme line at the early gastrula stage and in its descendants (Imai et al., 2004). Afterward, José-Edwards et al. (2011) identified a second Fos member expressed in the mesenchyme from the 110-cell stage and transiently in initial tailbud notochord cells.

Some pieces of information are available in the literature about *Ciona Jun* and *Fos* expression, but no data are reported on the other AP-1 members, and hypotheses on their functions have been formulated only on the basis of phylogenetic similarities with vertebrates (José-Edwards et al., 2011). *Jun* was identified in *Ciona* as a target of the Erk signaling, an evolutionarily conserved key pathway that regulates a wide variety of cellular processes, including proliferation, differentiation, apoptosis, and stress responses under both normal and neoplastic conditions (Kobayashi et al., 2018). All these functionalities are typically controlled by the AP-1 complex in other species, therefore suggesting that *Ciona* is a good model to investigate its functional and regulatory evolution (Amoutzias et al., 2006).

We have here further investigated the AP-1 complex in the ascidian *C. robusta* and analyzed its members' composition and their expression patterns during embryonic development. We, then, investigated in transgenic embryos the effects of *Jun* and *Fos* notochord-specific overexpression, giving, for the first time, an indication of their functional role in the structural organization of the cells.

MATERIALS AND METHODS

Animals

Adult specimens of *C. robusta* were collected in the Gulf of Taranto, Italy, by handpicking at low depth, and transported in seawater tanks to the facility of Stazione Zoologica Anton Dohrn (SZN). Animals were acclimatized at ∼18°C for 2–3 days in open system tanks and fed every day with a Shellfish Diet 1,800^{TM} Instant Algae®. Subsequently, they were exposed to continuous lighting for a few days to accumulate mature gametes and to prevent gamete spawning. Embryos were staged following the developmental timeline established by Hotta et al. (2007).

Paralog Prediction, Network Construction, and Functional Annotation

Jun and *Fos* transcription factor gene IDs (778972 and 778607, respectively) were retrieved from the NCBI Gene partition (Brown et al., 2015). The protein sequences of *C. robusta* assembly GCF_000224145.3 annotation release 104 were downloaded from the RefSeq partition at NCBI (O'Leary et al., 2016). An all-vs.-all similarity search of the entire *C. robusta* protein collection was performed using the BLASTp program of the BLAST package (Camacho et al., 2009), setting an *E*-value cutoff at 10^{-3}. Based on the detected similarity relationships, proteins were grouped in networks of paralogs according to the procedure described in Ambrosino et al. (2018) using the Network X package (Hagberg et al., 2008). A filtering step was introduced to select different *E*-value thresholds (e^{-10}, e^{-15}, and e^{-30}) that led to the construction of distinct networks of paralogs at different cutoff settings. A complete updating of the functional annotation of the *C. robusta* protein collection was performed with the software InterProScan (version 5.33) (Jones et al., 2014). UniProt proteins with the InterPro domains related to Jun, Fos, ATF3, and ATF2 were retrieved by keyword search performed on InterPro public database. Sequence similarity searches, using the detected *C. robusta* Jun, Fos, Atf3, and Atf2 protein members as probe queries, were performed by scanning the entire non-redundant protein database (nr) with the BLASTp web resource. Sequences from Tunicata, Cephalochordata, and Hemichordata, together with sequences from the echinoderm *Strongylocentrotus purpuratus* (sea urchin) as out-group, were retrieved. Moreover, sequences from the model species *Homo sapiens*, *Danio rerio*, and *Drosophila melanogaster* as reference species of Vertebrata and Protostomia were retrieved from the NCBI Gene partition. Multiple alignments of the retrieved sequences were performed with MAFFT software v7.397 (default parameters) (Katoh and Standley, 2013). Alignments were cleaned with TrimAl v1.4 (Capella-Gutiérrez et al., 2009) using a 0.25 gap threshold, 0.25 residue overlap threshold, and sequence overlap ranging from 50 to 75%. Maximum likelihood trees were constructed with FastTree v2.1.11 (Price et al., 2010) using default parameters. The visualization of the phylogenetic trees was performed with Interactive Tree of Life (iTOL) v4 (Letunic and Bork, 2019). Pieces of information about the domains contained in the multiple alignments were retrieved through the graphic summary of the NCBI BLAST web portal.

Isolation of AP-1 Member Transcripts

The full-length coding sequences (CDSs) of *Jun* and *Fos*, corresponding to the *Ciona* transcript models KH.C5.610.v1.A.SL1-1 and KH.C11.314.v2.A.ND1-1, respectively, were amplified by PCR using as template the cDNA derived from mRNA poly(A)$^{+}$ from a mix of *Ciona* embryos at tailbud stages (D'Aniello et al., 2011) and from the *Fos* clone of the Gateway-compatible Unigene collection (gift of Dr. Anna Di Gregorio). Similarly, the *Atf3* full-length cDNA corresponding to the transcript model KH.L5.5.v1.A.ND1-1 and its 3'-untranslated region (UTR) was amplified. Regarding *Maf* transcript (KH.C7.294.v1.A.SL1-1), only the sequence covering the last 1,432 bp of the CDS and the first 446 bp of its 3'-UTR was amplified. All the PCR fragments were cloned in the pCR® II-TOPO vector (Thermo Fisher Scientific) and sequenced. The primers are listed in **Supplementary Table 1**.

Construct Preparation

The p*Bra* > *Titf1* construct (Spagnuolo and Di Lauro, 2002) was used to prepare the constructs used for the co-electroporation experiments. It contains 790 bp of the promoter region of *Ciona Brachyury* gene (*Bra*) directing notochord-specific patterns of expression in both primary and secondary lineages (Corbo et al., 1997), the *Ciona* endodermal marker *Titf1* gene, and the SV40 poly-A signal. To generate *Bra* > *Jun* and *Bra* > *Fos*, the *Titf1* gene was removed, by digesting the vector with *Bam*HI/*Mlu*I, and replaced with the *Jun* and *Fos* full-length cDNAs, amplified from their respective pCR® II-TOPO clones (see **Supplementary Table 1** for primer sequences).

To generate *Bra* > *RFP* and *Bra* > *LacZ*, the *Bra* promoter was excised from the p*Bra* > *Titf1* construct and inserted into the RFP Zeller vector (Zeller et al., 2006) or in the pBlueScript II KS 1,230 (gift of R. Krumlauf, Stowers Institute, Kansas City, United States), which contains the LacZ reporter gene and SV40 poly-A sequences.

Transgenesis in *Ciona robusta*

Constructs were electroporated into fertilized eggs as previously described (Locascio et al., 1999). Briefly, *Ciona* eggs were deprived of the chorion and of follicular cells and fertilized. The exceeding sperm was washed out through several passages in filtered seawater (FSW) then immediately transferred in a solution containing 0.77 M mannitol and plasmid DNA. Electroporations were performed in 0.4-cm cuvettes using a Bio-Rad Gene Pulser II at constant 50 V and 500–800 μ F.

The transgenic embryos were reared up to early neurula (stage 14), mid tailbud I (stage 21), initial tailbud (stage 17/18), early tailbud I (stage 19), and late tailbud I (stage 23) stages, fixed and treated for subsequent analyses [e.g., whole-mount *in situ* hybridization (WISH), phalloidin staining].

Transgenic experiments for morphological observations of the effects of 10, 20, and 40 μg of *Jun* and *Fos* overexpression in notochord precursors were carried out twice and in biological triplicate. Subsequent transgenic experiments were carried out in biological triplicate and repeated at least seven times. The reported percentages of positive signals correspond to the total of five independent electroporations.

Embryo Staining and Microscopy

LacZ Staining

Bra > LacZ transgene expression was visualized by detection of β-galactosidase activity as previously described (D'Aniello et al., 2011). After staining, embryos were washed in 1 × phosphate-buffered saline (PBS), and imaging capture was made with a Zeiss Axio Imager M1 microscope. A minimum of 100 embryos was analyzed in at least six different electroporations.

Phalloidin Staining

Transgenic embryos were fixed for 30 min in MEM-PFA [4% paraformaldehyde (PFA), 0.1 M MOPS pH 7.4, 0.5 M NaCl, 1 mM EGTA pH 8.0, 2 mM $MgSO_4$ in H_2O], then washed four times for 20 min in PBT (1 × PBS-0.01% Triton X-100). They were then incubated at room temperature (RT) in PBT2 + 1:100 Alexa Fluor 635 phalloidin (Thermo Fisher Scientific) for 4 h and rinsed for 5 min in PBT. Afterward, embryos were washed four times for 20 min, alternating PBT and 1 × PBS. Embryos were imaged with a Leica TCS SP8X confocal laser scanning using the HC PL APO CS2 40 × /0.85 dry objective.

Whole-Mount *in situ* Hybridization

Single WISH and double WISH were carried out as previously described (Christiaen et al., 2009). Briefly, wild-type embryos for *in situ* experiments were obtained by *in vitro* fertilization and fixed at the desired stages in 4% PFA, 0.1 M MOPS pH 7.5, and 0.5 M NaCl at 4°C overnight.

The *Fos* probe was generated from the 63M13 clone found in the *Ciona intestinalis* Gateway-compatible Unigene collection (Beckman Coulter Genomics, Grenoble, France) (gift from Dr. A. Di Gregorio, New York University College of Dentistry, NY, United States). For the synthesis of *Jun*, *Atf3*, and *Maf* probes, their corresponding TOPO-TA clones were used.

The characterization of the mutant embryos by WISH was performed using notochord-, endoderm-, muscle-, and mesenchyme-specific genes chosen for their definite expression pattern on the base of the data reported in ANISEED at the initial tailbud stage (Brozovic et al., 2018). In particular, the following gene models were considered: KH.C3.225 and KH.C13.35 [human Fibrillinn (FBN1) and Chondromodulin (CNMD) orthologs, respectively) for the notochord, KH.L141.45 for the endoderm, KH.C8.859 [human myosin light chain 2 (MYL2) ortholog] for the muscles, and KH.C1.222 and KH.C5.202 [human aldo-keto reductase family 1 member C1 (AKR1C1) and TAL bHLH transcription factor 2 (TAL2) orthologs, respectively] for the mesenchyme. The clones corresponding to all these genes were found in the *Ciona* Gene Collection (Satou et al., 2002) with, respectively, the following IDs: R1CiGC02k18, R1CiGC05o23, R1CiGC02p08, R1CiGC24p06, R1CiGC01b11, and R1CiGC11k01.

A Zeiss Axio Imager M1 was used for embryo image capture. Pictures were edited with Adobe Photoshop CS6, and adjustments, where applied, were only for clarity without affecting any essential part of the image.

RESULTS

Searching for AP-1 Gene Members in *Ciona robusta* Genome

The entire protein collection of *C. robusta* has undergone a comprehensive functional annotation to better characterize all the genes classified as bZIP members related to *Jun* and *Fos* (hereafter called AP-1 members). An all-against-all protein sequence similarity search was performed on the entire protein collection of *C. robusta.* This analysis led to the construction of networks of paralogs that were calculated based on three different *e*-value thresholds with the aim to identify sets of paralog proteins based on different similarity cutoffs. The use of more stringent *e*-value cutoffs, indeed, defines a smaller number of paralogy relationships between proteins, obtaining a larger number of networks, when compared to paralog detections that employ less stringent *e*-value thresholds. This permits to define subgroupings of more similar proteins within groups of less similar protein-encoding genes (i.e., in networks of computationally predicted paralogs defined at less stringent *e*-value thresholds).

The results of the performed functional annotation analysis were considered in light of the networks of paralog organization in order to identify all those containing the members of the AP-1 family. As represented in **Figure 1**, members of Fos, Atf3, and Maf (a gene named LOC778711 that in this study we called *Maf* because it encodes a protein with a bZIP Maf motif) are grouped within the same network of paralogs when using the less stringent *e*-value cutoff (e^{-10}), while Jun and Atf2 members show a distinct organization. Moving to a more stringent cutoff (*e*-value at e^{-15}), the Maf member splits from the Fos/Atf3 group. The use of the more stringent *e*-value threshold (e^{-30}), finally, discriminates also Fos from Atf3 members.

A search of all the detected *C. robusta* AP-1 family members in the Gene Partition at the NCBI site revealed that Fos, Atf3, Atf2, Jun, and Maf proteins are isoforms each encoded by a single gene. In particular, we detected four protein isoforms encoded by the *Fos* gene, four isoforms encoded by *Atf3*, two isoforms encoded by *Atf2*, and single proteins encoded by *Jun* and *Maf*, respectively (**Supplementary Table 2**). A similar search on the considered model species revealed the presence of multiple copies of *Jun* and *Fos* genes in Vertebrata, i.e., three genes related to *Jun* family (*Jun*, *JunB*, and *JunD*) and four genes related to Fos family (*Fos*, *FosB*, *FosL1*, and *FosL2*) in *H. sapiens*; six *Jun* genes (*Jun*, *JunBa*, *JunBb*, *JunD*, *JunDP2a*, and *JunDP2b*) and seven *Fos* genes (*FosAa*, *FosAb*, *FosB*, *FosL1A*, *FosL1B*, *FosL2*, and *FosL2L*) in *D. rerio*. In the considered model from the Protostomia *D. melanogaster*, all the considered AP-1 family members are encoded by a single gene. Details about the number of protein isoforms encoded by these genes are reported in **Supplementary Table 2**.

To verify the consistency of the results obtained in *C. robusta* and to provide a rough overview of what is available about the AP-1 family members on the reference protein databases, we interrogated InterPro web portal to retrieve all the Uniprot proteins having functional annotation related to Jun, Fos/Atf3/Maf, and Atf2 in all chordate species. By this approach, we obtained 565 different chordate species having 1,993 Jun-related proteins, six of which were protein isoforms from tunicates; 572 different chordate species having 5,505 Fos/Atf3/Maf-related proteins, 10 of which were from tunicates; 511 chordate species having 1,007 Atf2-related protein entries, three of which were from tunicates. All the results are summarized in **Supplementary Table 3**, and they highlight the widespread presence of AP-1 family proteins among chordates.

The selection of the sequences to be included in the phylogenetic analysis was done by performing sequence similarity searches of the reference NCBI gene sequences of *C. robusta Jun*, *Fos*, *Atf3*, *Atf2*, and *Maf* protein members as probe queries against the entire non-redundant protein database (nr). The probe queries included the reference or the longest isoform for each gene. The best hit for each detected species among the Tunicata, Cephalochordata, and Hemichordata, together with the echinoderm *S. purpuratus* as out-group, was retrieved. The sequences from the model species *H. sapiens*, *D. rerio*, and *D. melanogaster* were selected as reference species of Vertebrata and Protostomia. Most importantly, no significant hit related to the *Maf* probe sequence was detected in the other species, probably revealing a species-specific protein.

The multiple alignments performed with the collected Jun sequences (**Supplementary Figure 1**) highlighted the presence of the conserved domains "bzip_Jun" (cd14696) and "Jun-like transcription factor" (pfam03957). Accordingly, the retrieved Fos sequences (**Supplementary Figure 2**) showed the presence of the conserved domain "bZIP_Fos" (cd14721), and the ATF3 and ATF2 sequences (**Supplementary Figures 3**, **4**) highlighted the presence of the conserved domains "bZIP_ATF3" (cd14722) and "bZIP_ATF2" (cd14687), respectively. As expected, all proteins grouped in each alignment showed the presence of the conserved domains "bZIP transcription factor" (pfam00170) and "basic region leucine zipper" (smart00338).

The unrooted phylogenetic tree generated from the collected Jun sequences (**Supplementary Figure 5**) highlighted the presence of a certain sequence conservation under each taxon. In particular, the following can be easily recognized: a three-leaf branch for Tunicata followed by a cluster of seven leaves for Vertebrata; a three-leaf branch including Cephalochordata and single leaves of Echinodermata, Hemichordata, and Protostomia sequences. The only exception to the clusterization of sequences from the same taxon is represented by sequences from the tunicate *Oikopleura dioca* and zebrafish (*D. rerio*), i.e., the two Jun dimerization proteins 2 (JUNdp2a and JUNdp2b). The unrooted phylogenetic tree generated from the collected Fos sequences (**Supplementary Figure 6**) showed a similar behavior. Again, we observed a Tunicata branch (with two leaves) related in the phylogenetic tree to a cluster of Vertebrata sequences including multiple copies from human and zebrafish and to a three-leaf branch of Cephalochordata. Also in this case, we can observe for *Oikopleura* and for the FOSL2L gene from zebrafish a position in the tree that is distant from other tunicate or vertebrate sequences and more closely related to leaves of Hemichordata, Echinodermata, and Protostomia (**Supplementary Figure 6**).

Since we were not able to detect in other species robust homologs of the *C. robusta Maf* gene, we decided to include the sequence of its encoded protein in the phylogenetic analysis related to another member of the AP-1 family. As inferred by our paralog analysis performed on the entire *Ciona* protein collection, *Maf* resulted in the same network of paralogs of *Fos*- and *Atf3*-related protein isoforms when using the less stringent *e*-value cutoff (**Figure 1**). We tried to perform a new phylogenetic tree including *Maf* protein sequence among *Atf3* and *Fos* sequences. Only the set of *Atf3* sequences enabled the conservation of *Maf* sequence in the multiple alignment after the trimming step. This was probably due to the marked difference between the lengths of *Maf* and *Fos* sequences. The tree generated from the retrieved *Atf3* sequences, including the *Maf* sequence from *C. robusta* (**Supplementary Figure 7**), showed the presence of three main ATF3 branches with Tunicata ATF3 sequences rather distant from the rest of the considered sequences that are, respectively a two-leaf branch of Vertebrata close to a two-leaf branch of

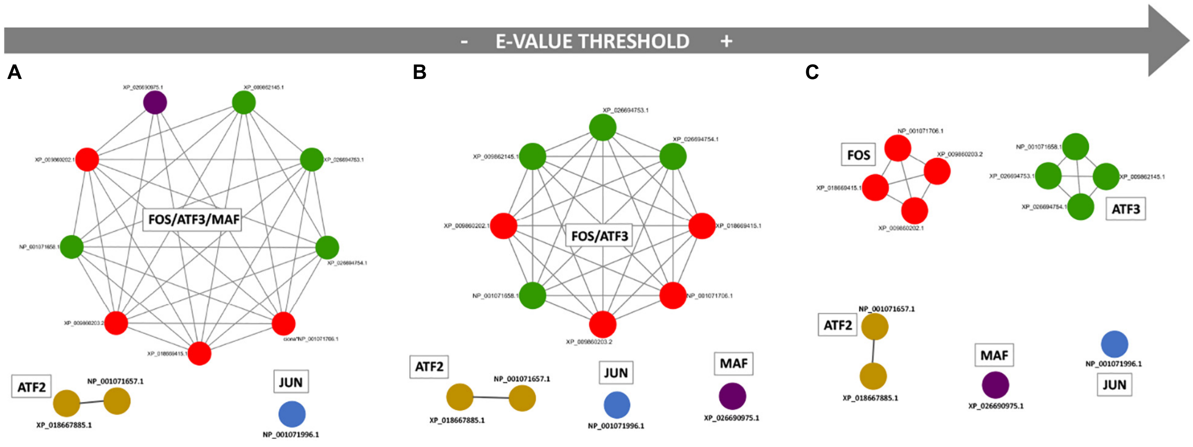

- E-VALUE THRESHOLD +
A
XP_026690975.1
XP_009862145.1
XP_009860202.1
XP_026694753.1
FOS/ATF3/MAF
NP_001071658.1
XP_026694754.1
XP_009860203.2
ciona*NP_001071706.1
XP_018669415.1
ATF2
NP_001071657.1
XP_018667885.1
JUN
NP_001071996.1
B
XP_026694753.1
XP_009862145.1
XP_026694754.1
XP_009860202.1
FOS/ATF3
XP_018669415.1
NP_001071658.1
NP_001071706.1
XP_009860203.2
ATF2
NP_001071657.1
XP_018667885.1
JUN
NP_001071996.1
MAF
XP_026690975.1
C
NP_001071706.1
FOS
XP_009860203.2
XP_018669415.1
XP_009860202.1
NP_001071658.1
XP_026694753.1
XP_009862145.1
XP_026694754.1
ATF3
NP_001071657.1
ATF2
XP_018667885.1
MAF
XP_026690975.1
NP_001071996.1
JUN

FIGURE 1 | *Ciona robusta* bZIP members related to Jun and Fos when similarity threshold changes. Networks of paralogs were detected at different *e*-value thresholds: **(A)** e^{-10}; **(B)** e^{-15}; **(C)** e^{-30}. Each node in the network represents a protein, and each gray edge represents a paralogy relationship. FOS member isoforms are shown in red, ATF3 in green, ATF2 in light brown, MAF in purple, JUN in blue.

Cephalochordata and a second branch with leaves of sequences from Protostomia, Echinodermata, and Hemichordata. The position of *C. robusta Maf* sequence between Cephalochordata and *D. melanogaster* in the phylogenetic tree probably indicates an ancestral separation of this sequence that soon diverged from the other related sequences of the *ATF3* family. The tree generated from the collected *Atf2* sequences (**Supplementary Figure 8**) showed a three-leaf branch of tunicates close to a two-leaf branch including *Oikopleura dioica* and *D. melanogaster*. More distant in the tree, we can observe the presence of *Atf2* sequences from Hemichordata and Echinodermata, followed by the considered Cephalochordata and Vertebrata taxa.

The phylogenetic analysis revealed in all the four considered classes an ancestor behavior of AP-1 members of Tunicata subphylum with respect to the other considered sequences. An exception to this behavior is represented by the sequences from *Oikopleura*, a tunicate with peculiar genomic features that make it closer to Protostomia in the phylogenetic trees rather than other tunicates (**Supplementary Figures 5**, **6**, **8**). The performed analyses highlighted, moreover, the presence of single gene copies in tunicates opposed to gene amplification of *Jun* and *Fos* members in Vertebrata subphylum (**Supplementary Figures 5**, **6**). Zebrafish in particular has undergone extra-duplication events of its genome, as resulted from the number of copies of Jun and Fos genes (**Supplementary Table 2**).

Expression of AP-1 Members During Embryonic Development

Few data on *Ciona* Fos and Jun factors are currently available in the literature. We hence proceeded with a comparative analysis of the expression profiles of the AP-1 gene members during the embryonic development at neurula, early tailbud, and late tailbud stages (stages 15, 19, and 23, respectively). We assayed by WISH *Jun*, *Fos*, *Atf3*, and *Maf* expression, except for *Atf2/7* of which we were unable to amplify a valid riboprobe. As shown in **Figure 2**, we obtained very similar expression profiles in specific mesenchymal populations. As already reported in ANISEED database[1] and by Imai et al. (2004), *Jun* is expressed in a pair of B-cell mesenchymal cells from neurula to tailbud stages (**Figures 2D–F**); the same narrow signal is detectable for *Atf3* and *Maf* (**Figures 2J–O**), although *Atf3* appears later at tailbud stage 19. According to single-cell transcriptomic analyses (Cao et al., 2019), *Fos* staining shows a relatively stronger and broader signal in the mesenchymal cells at neurula and tailbud stages (**Figures 2G,H**) with a very transient signal in notochord precursors at the neurula stage (José-Edwards et al., 2011; data not shown). A novel, not previously reported, and symmetric *Fos* signal is clearly visible in the anterior part of the nervous system at late tailbud I stage (**Figure 2I**).

Functional Analysis of *Jun* and *Fos* Transcription Factors

Since our WISH experiments show an overlapping profile of the ascidian AP-1 gene expression, their heterodimerization faculty might be postulated. To corroborate this hypothesis, we focused our functional analysis on Jun and Fos members. The Jun proteins exist as homodimers and heterodimers, while the Fos proteins are historically considered able to form stable heterodimers with Jun proteins, although it has been demonstrated that, in peculiar conditions, they can also homodimerize. We, hence, assayed their activity by overexpression of their CDSs during notochord development under the control of the *Brachyury* promoter (Bra > Jun, Bra > Fos constructs). This tissue-specific expression would permit to avoid possible interference by other members of this family, and it is already reported in other species as a territory with functional significance for Jun and Fos activity.

Different concentrations of Bra > Jun and Bra > Fos constructs (10, 20, and 40 μg) alone or in combination were assayed by electroporation. The transgenic embryos were reared up to the mid tailbud I (stage 21), following the notochord cell intercalation, and Jun and Fos overexpression was evaluated by WISH. Compared to the control (*Bra > LacZ*), the overexpression in the notochord cells of *Jun* and *Fos* alone apparently did not influence the proper development of the embryos (**Figures 3A–C**).

Conversely, *Jun* and *Fos* co-electroporation resulted in a scale of abnormal phenotypes proportional to the amount of DNA used (**Figures 3D–I**) and ranging from "wt-like" in which all the structures of the embryo were perfectly recognizable (**Figure 3D**) to a completely aberrant one (**Figures 3H,I**).

Intra-electroporation variability was probably also linked to the stochastic nature of the method, where the variable amount of DNA that enters the fertilized eggs during electroporation is random and unpredictable. Within this scenario, as the quantity of electroporated DNA increases, the percentage of severely defective embryos increases too (**Table 1**). Indeed, WISH analyses on these embryos revealed that *Jun* and *Fos* expression in notochord cells was barely detectable in the wt-like tailbuds. Hence, depending on the level of overexpression induced in the notochord cells, three groups of embryonic phenotypes have been distinguished (**Table 1**): a first, almost normal, group of embryos with a nearly perfect development (**Figure 3D**); a second, moderately defective, group specifically displaying various tail alterations (**Figures 3E,F**); and finally, a third, severely defective, group exhibiting a deep displacing of the notochord and severe alterations throughout the whole embryo (**Figures 3G–I** and **Table 1**). It is worth noting that the development of Bra > *Jun* + Bra > *Fos* embryos proceeds as in control embryos up to the neurula stage, although *Brachyury* promoter starts to be active at the 64-cell stage (Corbo et al., 1997). Therefore, only the formation of tailbud embryos is affected, that is, in correspondence to notochord elongation and cell intercalation stage.

Together, these results confirm the ability of the Jun/Fos heterodimeric complex to influence notochord development and seem to indicate that, in *Ciona*, this complex can influence the transcriptional regulation of developmental genes.

---

[1] http://www.aniseed.cnrs.fr/

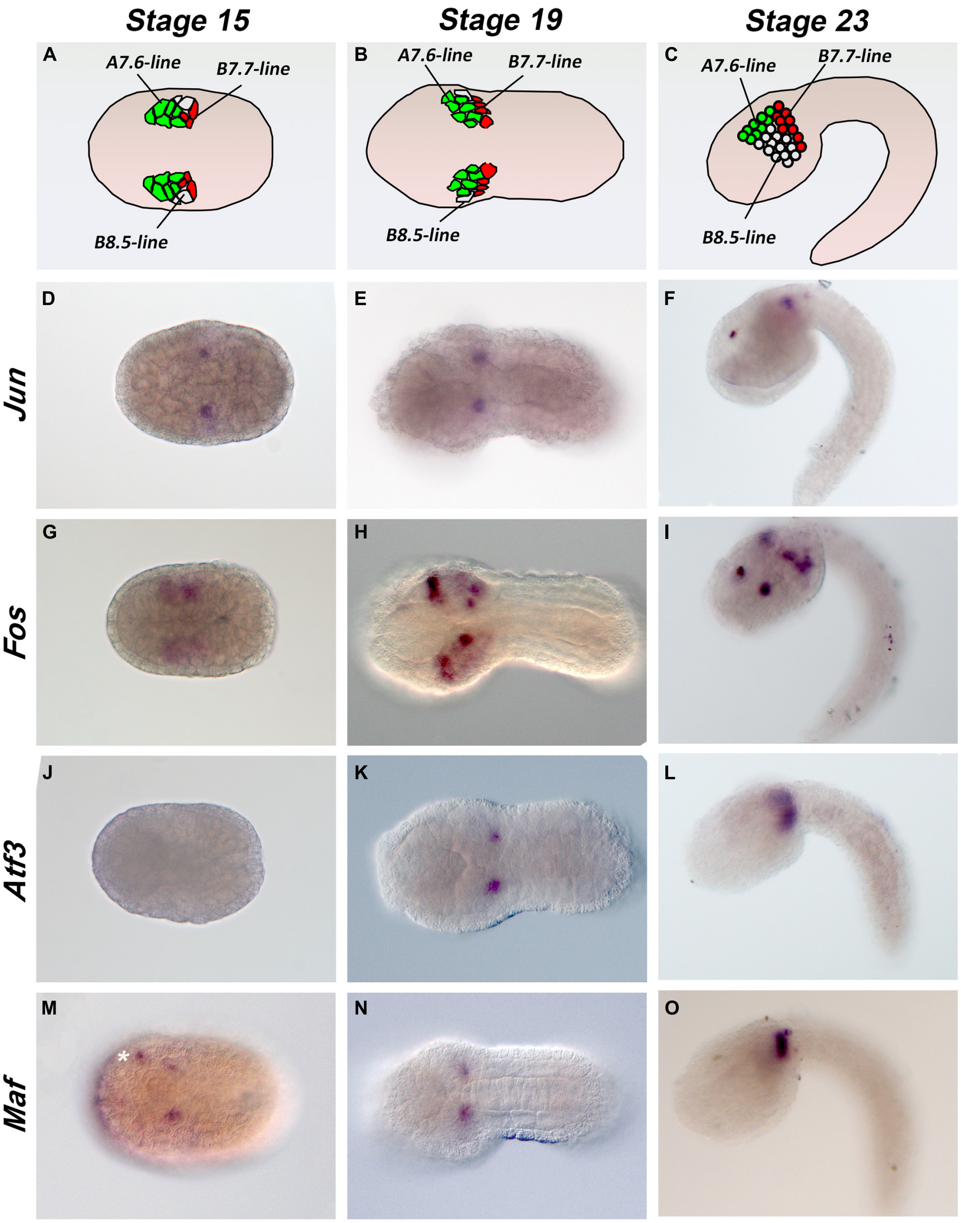


FIGURE 2 | Expression patterns of *Ciona* AP-1 transcription factors during embryonic development. **(A–C)** Schematic representation of mesenchymal cells in stage 15, 19, and 23 embryos. **(D–O)** Whole-mount *Ciona* embryos hybridized *in situ* with antisense RNA probes against the *Jun* **(D–F)**, *Fos* **(G–I)**, *Atf3* **(J–L)**, and *Maf* **(M–O)** genes, as indicated on the left side of each row. Developmental stages are indicated at the top of each column: stage 15, early neurula; stage 19, early tailbud I; stage 23, late tailbud I. White asterisk indicates an unspecific signal. Stage 15 and 19 embryos are in a dorsal view, and stage 23 embryos are in a lateral view; anterior is on the left.

Characterization of Bra > *Jun* and Bra > *Fos* Transgenic Tailbud Embryos

Bra > *Jun* and Bra > *Fos* overexpression in the notochord precursor cells induces progressively more severe alterations with increasing Jun and Fos concentrations. We, then, evaluated the differentiation levels of the various embryonic tissues on transgenic tailbud embryos (stage 17/18) electroporated with 10 μg of Bra > Jun and Bra > Fos DNA and split out in three separate groups according to their apparent phenotype severity. We assayed by WISH experiments the expression of muscle-, mesenchyme-, endoderm-, and notochord-specific markers selected from the ANISEED database for their definite and specific expression pattern (Brozovic et al., 2018).

Muscles and Mesenchyme

The overexpression of *Jun* and *Fos* in notochord cells appeared not to affect the differentiation of muscle and mesenchymal cells, as far as could be seen by examination, in the manipulated

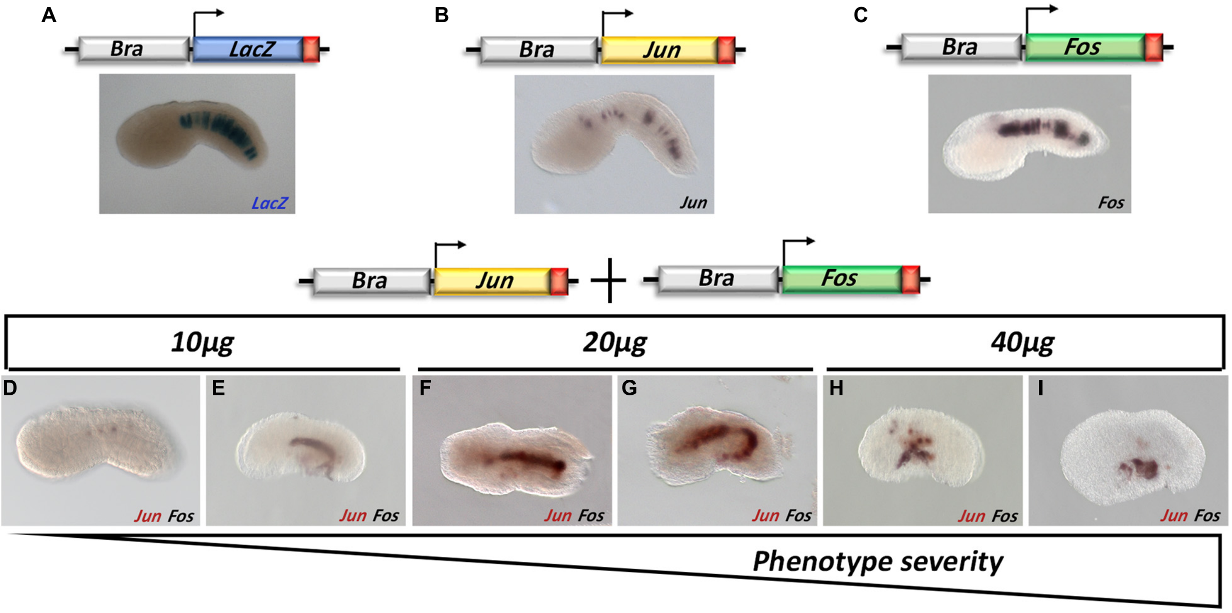


FIGURE 3 | Functional analysis of *Jun* and *Fos* during *Ciona robusta* development. **(A)** Brachyury promoter (Bra) specifically drives the expression of a transgene in the notochord cells; fertilized eggs electroporated with Bra > *LacZ* **(A)**, Bra > *Jun* **(B)**, and Bra > *Fos* **(C)** and reared up to the mid tailbud I stage (S21) were checked for the transgene expression through β-galactosidase assay staining (Bra > *LacZ*) or whole-mount *in situ* hybridization (WISH) using *Jun* (brown staining)- and *Fos* (black staining)-specific probes. **(D–I)** *Jun* and *Fos* double WISH on tailbud embryos co-electroporated with Bra > *Jun* and Bra > *Fos*. Increasing amounts of Bra > *Jun* and Bra > *Fos* go at the same pace with the growing severity of the phenotype; double WISH experiments show that the transgenes are less expressed in the most wt-like embryos **(D)**. For all embryos, the anterior is on the left and the dorsal on the top.

TABLE 1 | Percentages of phenotypic alterations as a function of the amount of Bra > Jun and Bra > Fos electroporated DNA.

Bra > Jun + Bra > Fos	**Normal**	**Moderately defective**	**Severely defective**
10 µg	30%	60%	10%
20 µg	20%	35%	45%
40 µg	2%	38%	60%

tailbud embryos of the global expression of the muscle-specific gene, *Myosin Light Chain 2* (*Myl2*) (Macera et al., 1992; **Figures 4A,A'**), and of two mesenchyme-specific genes, *Aldo-Keto Reductase Family 1 Member C1* (*Akr1c1*) (Ciaccio and Tew, 1994; **Figures 4B,B'**) and *TAL BHLH Transcription Factor 2* (*Tal2*) (Xia et al., 1991; **Figures 4C,C'**).

Endoderm

To assess the correct endoderm differentiation, control and co-electroporated tailbuds were tested for the expression of two endodermal markers, the alkaline phosphatase (*AP*) and a gene, not yet annotated, identified with the ID KH2012:KH.L141.45, that here we refer to as *02p08*. The first was evaluated through an AP activity assay (**Figure 4D**), the other one by WISH (**Figure 4E**). These tests showed that the Bra-directed overexpression of *Jun* and *Fos* did not affect the differentiation and localization of the endodermal cells.

Bra > *Jun* and Bra > *Fos* in Notochord Organization

The effects of the overexpression of *Jun* and *Fos* on the differentiation of notochord cells were first explored by WISH of the expression of two different markers exclusively expressed in this tissue: *Fibrillin 1* (*Fbn1*) (Sakai et al., 1986) and *Chondromodulin* (*Cnmd*) (Shukunami and Hiraki, 2001; **Figures 4G,H**). Furthermore, the activity of the *Bra* promoter was also assayed by tri-electroporation with Bra > *Jun* and Bra > *Fos* of the Bra > *RFP* construct containing the red fluorescent protein (RFP) downstream of the *Bra* promoter (Corbo et al., 1997; **Figure 4F**). From these analyses emerged that the expression of the examined notochord markers shows the same pattern as in the controls independently from their phenotype severity and that notochord cell differentiation is not affected. Regardless of the grade of progressive disorganization in the different phenotypes, by terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assay, we also observed that notochord alterations were not caused by the induction of apoptosis (data not shown) and that the number of notochord cells was comparable to that of the controls. These results suggest that the Jun and Fos could be involved in the structural organization of the notochord but not in its specification, proliferation, and differentiation.

To deeply explore the effects of Bra > *Jun* and Bra > *Fos* expression on the notochord morphology, transgenic embryos were left to develop until the late tailbud I stage (stage 23), when the notochord cell intercalation is completed, but before lumen

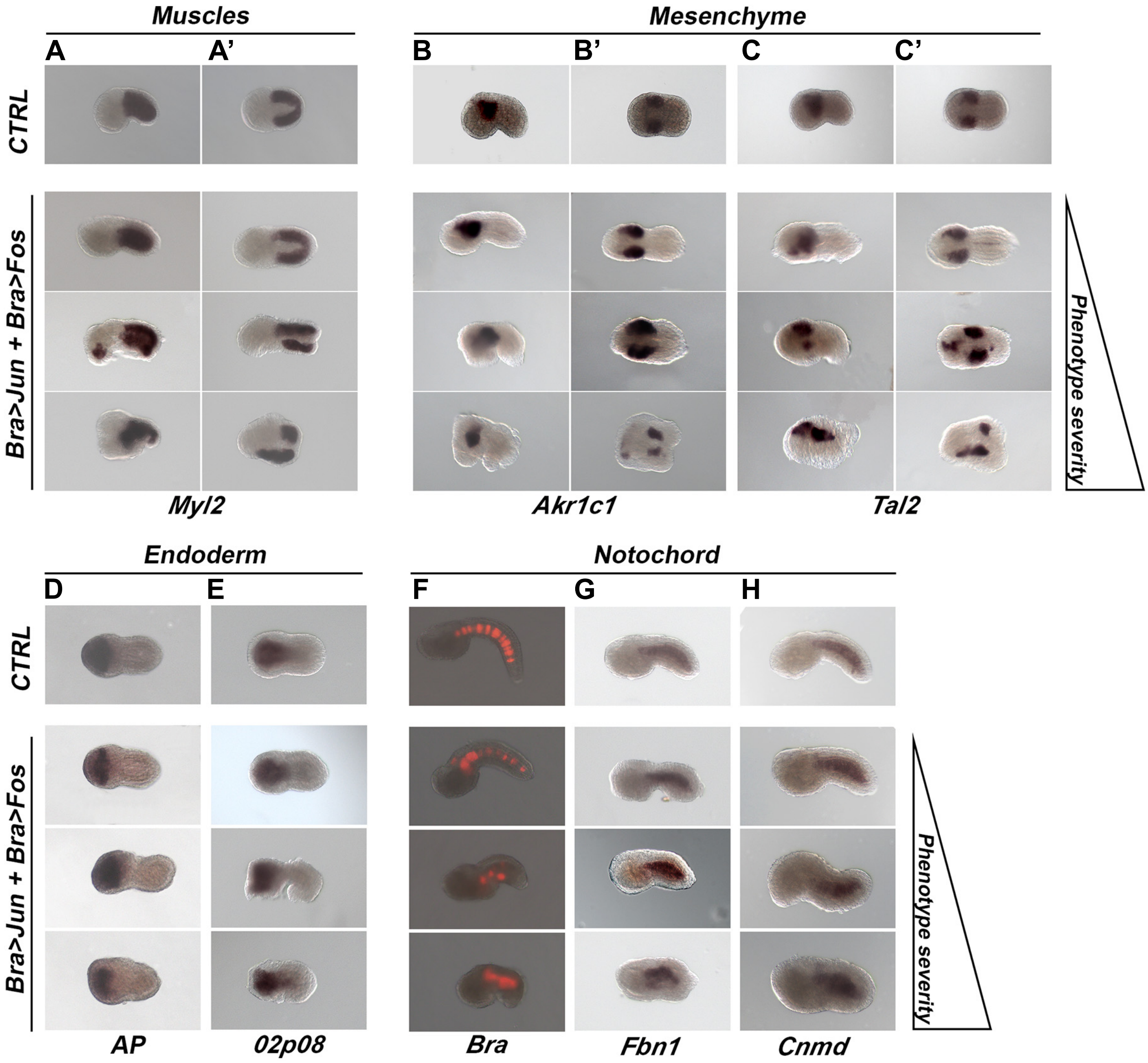


FIGURE 4 | Tissue differentiation in transgenic embryos (stage 17/18) co-electroporated with 10 μg of Bra > *Jun* and Bra > *Fos* constructs. **(A,A')** Muscle differentiation; whole-mount *in situ* hybridization (WISH) experiments with the muscle marker gene *Myosin Light Chain 2* (*Myl2*). **(B–C')** Mesenchyme differentiation; WISH experiments with mesenchyme markers **(B,B')** *Aldo-Keto Reductase Family 1 Member C1* (*Akr1c1*); **(C,C')** *TAL BHLH Transcription Factor 2* (*Tal2*). **(D,E)** Endoderm differentiation. **(D)** Alkaline phosphatase (AP) assay to detect the endogenous AP enzymatic activity. **(E)** WISH experiment with the endoderm-specific gene KH2012:KH.L141.45. **(F–H)** Notochord differentiation. **(F)** Fluorescent images showing reporter gene [red fluorescent protein (RFP)] expression under the control of the *Brachyury* (*Bra*) promoter in embryos tri-electroporated with 10 μg of Bra > *Jun*, Bra > *Fos*, and Bra > RFP. **(G,H)** WISH experiments with the notochord markers **(G)** *Fibrillin 1* (*Fbn1*) and **(H)** *Chondromodulin* (*Cnmd*). CTRL, Bra > RFP control embryos; Bra > Jun + Bra > Fos, co-electroporated embryos. In each column, embryos have been placed according to the progressive severity of the phenotype: normal (top), mild (middle), and severe (bottom) alterations. For all embryos, the anterior is on the left. **(A–C,F–H)** Lateral view; **(A',B'C',D,E)**, dorsal view.

formation (**Figure 5A**), and then, the transgenes were observed under a fluorescence microscope (**Figures 5B–F**). For this analysis, we again used 10 μg of Bra > *Jun* and Bra > *Fos* alone or in combination and added in all assays the Bra > RFP construct as an internal control to selectively evaluate the correct formation of electroporated notochord cells. Furthermore, phalloidin staining of the embryos was used to define cell membrane profiles of the notochord. Number counting of RFP-positive embryos evidenced a notable decrease, almost by half, of developed embryos among those electroporated with Jun and Fos together (**Supplementary Table 4**). As annotated in **Table 2**, the electroporation efficiencies, evaluated based on Jun/Fos embryo RFP expression, were comparable in all samples, ranging from 83.15% of the control Bra > RFP embryos to 92.31% of the tri-electroporated. Among the fully developed embryos, we then evaluated notochord impairment and considered normal only the notochords with a perfect alignment of all cells. Following electroporation of the Bra > RFP construct, control larvae displayed the wild-type

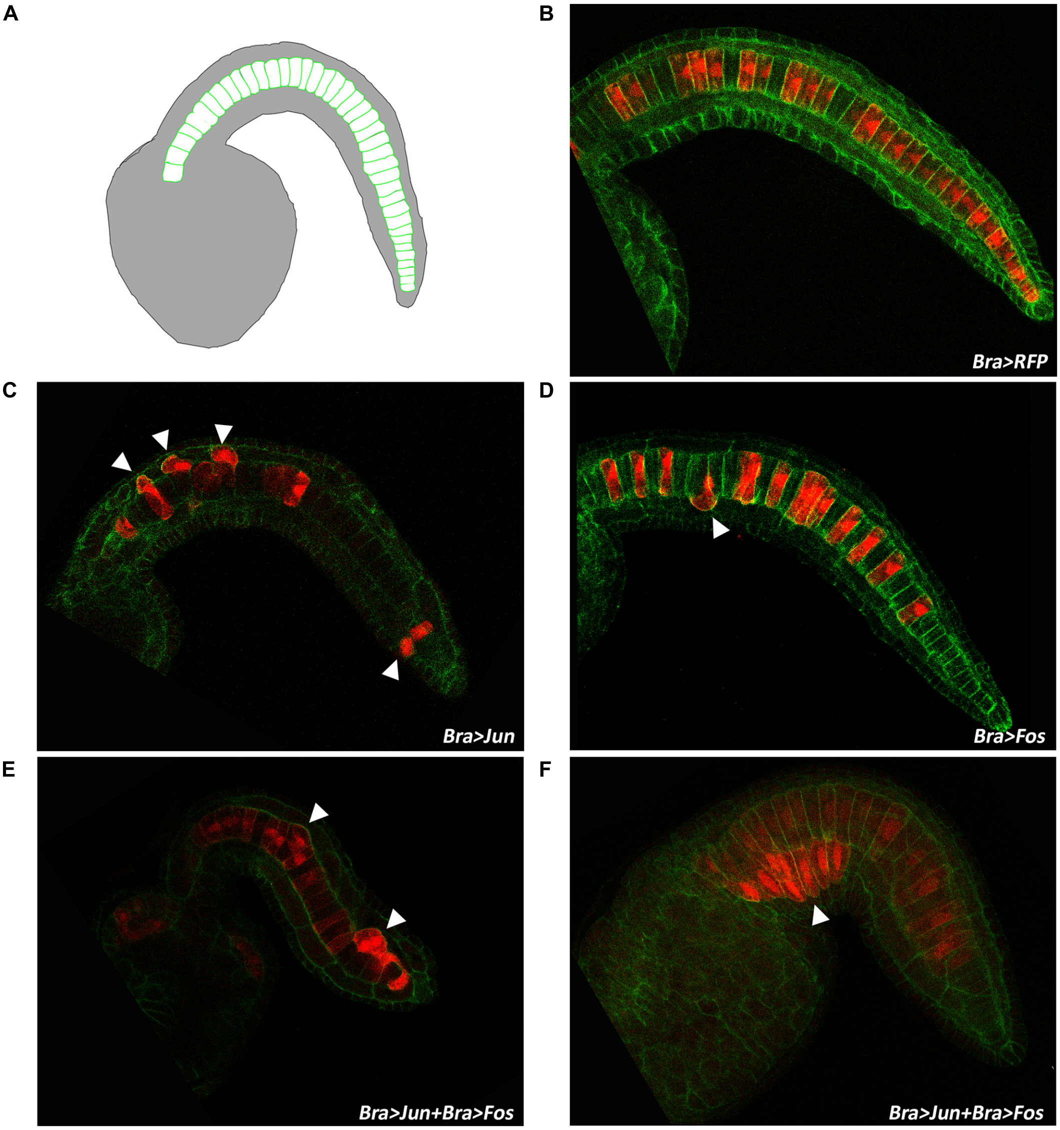


FIGURE 5 | Intercalation defects in tailbud embryos overexpressing *Jun* and *Fos*. **(A)** Schematic representation of notochord cell organization in the late tailbud embryo (stage 23). **(B)** Confocal microscopy of a stage 23 control embryo electroporated with Bra > RFP. **(C,D)** Stage 23 embryos co-electroporated with Bra > RFP and either Bra > Jun or Bra > Fos, respectively. **(E,F)** Embryos at the same stage, co-electroporated with Bra > RFP and Bra > Jun plus Bra > Fos showing altered organization of notochord cells along the tail. White arrowheads evidence misaligned cells. Red, red fluorescent protein (RFP) staining; green, phalloidin staining.

notochord phenotype (**Figures 5A,B**). In *Jun* and *Fos* transgenic embryos, we observed two types of statistically significant alterations in notochord cells. In particular, about 11% of Bra > Fos and 23% of Bra > Jun embryos displayed single notochord cell misalignments (**Figures 5C,D**). This incorrect positioning of single notochord cells increases to 30% in Bra > Jun plus Bra > Fos co-electroporated embryos (**Table 2**). Notably, a second more severe misalignment of notochord cells was mostly specific for *Jun* and *Fos* co-electroporated samples. In 29% of *Jun/Fos* embryos, we observed a displacement of groups of notochord cells (**Figures 5E,F**). This wider notochord disorganization has always been identified when no less than 4–5 consecutive cells were electroporated as demonstrated by their RFP-positive signal (**Figures 5E,F**). Conversely, we never observed single or in blocks misalignments in the notochord when there were no consecutive cells electroporated. Cell misalignment equally involved anterior and posterior notochord cells (**Figures 5E,F**), thus revealing an equal impact of Jun and Fos on notochord cells deriving from primary and secondary lineages, respectively.

The 23% of single-cell alterations in embryos electroporated with *Jun* might be easily explained if we consider the homodimerization ability of this TF. Conversely, 11% of misalignments in embryos overexpressing *Fos* could be

TABLE 2 | Percentages of notochord alterations observed in Bra > Jun and Bra > Fos transgenic embryos.

Constructs	Electroporation efficiency %	Fully developed embryos N.	Notochord phenotypes		
			Normal (%)	Single cell alteration (%)	Blocks of cells alteration (%)
BRA > *RFP*	83.15%	178	67.98%	2.70%	0.68%
BRA > *Fos*	91.35%	104	72.11%	11.58%	1.05%
BRA > *Jun*	88.55%	131	62.93%	23.27%	6.03%
BRA > *Jun* + BRA > *Fos*	92.31%	104	32.29%	30.21%	29.17%

presumably the consequence of the formation of Fos homodimers or heterodimers with unknown factor/s. Whichever is the most correct hypothesis, these complexes seem in any case less effective than Jun/Fos combined activity, and further studies will be necessary to verify if the phenotype derives from the formation of Jun–Fos heterodimers.

In particular, the observed partial convergence of notochord cells to form two adjacent and opposing rows (**Figures 4**, **5E,F**) is very similar to the knockdown phenotype of *Fibronectin* (*Fn*) gene (Segade et al., 2016). *Fn* gene is required for proper notochord intercalation in *Ciona* and, as well as in *Jun/Fos* transgenic embryos, its functional perturbation does not disrupt sheath integrity and defective cells do not escape into adjoining tissues (**Figures 5E,F**). Given this similarity in phenotypes, to assess if *Jun* and *Fos* co-expression can affect notochord intercalation through the negative regulation of *Fn*, we analyzed the expression of this gene by WISH on transgenic embryos (**Supplementary Figure 9**). The result was that *Fn* expression does not change in any of the embryos with *Jun* and *Fos* ectopic expression alone or in combination (**Supplementary Figures 9B–E**). *Fn* expression remains unaltered even in severely defective phenotypes in which the notochord cells did not intercalate and are still placed in two separate rows (**Supplementary Figure 9E**).

Taken together, these data suggest that the overexpression of *Jun* and *Fos* affects notochord cell intercalation but not through *Fn* regulation.

DISCUSSION

Despite all the existing data, it is difficult to draw a clear picture about the general physiological role for the AP-1 transcriptional complex in cellular processes and in the embryonic development. Based on the literature, it is quite obvious that the abundance of different AP-1 members within given cell types, as well as cell lineages, differentiation stage, microenvironment, and type of stimulus, has a large impact on how AP-1 modulates the determination of cells to proliferate, differentiate, or die by apoptosis.

To better characterize in *C. robusta* the predicted paralogs of the AP-1 family members and their evolutionary relationships, we searched for sequence similarities in the NCBI and Uniprot databases and performed phylogenetic analyses of the retrieved sequences.

The genomic locus assignment of each protein sequence revealed the presence of single gene copies for each *Ciona* AP-1 factor. In particular, single Jun, Fos, Atf2, and Atf3 members and an additional *Ciona*-specific Maf member evidenced less redundancy for this bZIP family in ascidians and the presence of a highly conserved core of AP-1 members that was formed in the last common ancestor of chordates.

Individual members of each family may have different biological functions during the formation of the embryonic tissues. By comparative analysis of these gene expression profiles during *Ciona* development, it was evident that they all show a common and restricted pattern in the B7.7 mesenchyme line (**Figure 2**). In *C. robusta* and *savigny*, it was demonstrated that this line gives rise to tunic and blood cells (Tokuoka et al., 2005). In this regard, different *in vitro* and *in vivo* molecular genetic approaches have demonstrated the many roles of AP-1 (Fos/Jun) TF in the development and differentiation of hematopoietic precursor cells (reviewed in Liebermann et al., 1998). However, none of these studies has completely identified the whole molecular mechanisms responsible for these effects also because of the number of AP-1 family members and their highly variable dimeric composition. The presence of single gene copies for *Ciona Fos*, *Jun*, *AT3*, and *Maf* factors and their common expression in embryonic mesenchyme cells that will give rise to blood cells reveal interesting potentialities about the different combinations of dimeric factors they can form and their possible conserved function in hematopoiesis. In this regard, ascidians with their not duplicated number of AP-1 members may represent an excellent experimental system to explore the evolution of AP-1 mechanisms of action in hematopoiesis.

A wider *Fos* expression profile, not only in mesenchymal cells but also in notochord and neural precursors of the sensory vesicle, indicates that Fos may have functions that are Jun/Atf/Maf-independent and suggests that it may form homodimers or heterodimers with other than AP-1 factors (Szalóki et al., 2015).

Although we still do not know which are the partners of Fos in the notochord and its mechanism of action, our results after *Jun* and *Fos* ectopic expression in notochord cells evidenced that Fos and Jun alone can induce narrowed alterations limited to single cells, while only the combined presence of both factors induces structural disorganization and severe impairment of cellular intercalation. This result is perfectly in agreement with the expression of *Fos* only in the initial differentiation of the notochord cells and evidenced how important is its downregulation so that correct intercalation can occur. Further indirect evidence about the presence of AP-1 consensus motif in notochord *cis*-regulatory modules that are able to cooperate with Foxa2 in activating the expression of notochord-specific genes (José-Edwards et al., 2015) contributed to strengthening our conclusions. Our findings are in support of a conserved and yet unexplored genetic program regulated by Fos in notochord formation that additional information on *Ciona* AP-1 mechanisms of action could help clarify. Similarly, a fundamental role played by AP-1 complex was evidenced in vertebrates for axial skeletogenesis, for proper histo-architecture of the epidermis, and for mesenchymal–epithelial cross-talk in the skin (Angel and Szabowski, 2002; Behrens et al., 2003). Interestingly, Jun/Fos-induced alterations in notochord cells clearly resemble the phenotypes observed when the planar cell polarity (PCP) process is impaired. In vertebrates, the PCP pathway is important for developmental processes in a number of organs and tissues as in the control of convergence and extension during gastrulation and in the control of the cell behaviors that drive notochord intercalation (Wallingford et al., 2000; Topczewski et al., 2001). Interfacing the results obtained in ascidians and vertebrates on the PCP pathway and its connection with the AP-1 complex, interesting hypotheses of functional conservation can be applied. Key components mediating cell–cell communication in the PCP vertebrate signaling pathway include the Wnt non-canonical pathway, the transmembrane protein Frizzled (Fz), and the cytoplasmic proteins Prickle (Pk) and Disheveled (Dsh), which culminate in the activation of c-Jun N-terminal kinase (JNK) and of the AP-1 complex to promote actin cytoskeleton reorganization and cellular movements. Comparatively, the ascidian Wnt5, Prickle, and Dsh mutants result in loss of PCP signaling and of notochord cell intercalation in the mediolateral axis (Keys et al., 2002; Jiang et al., 2005; Niwano et al., 2009). In light of this conserved PCP pathway, it will be highly interesting to investigate the involvement of the AP-1 transcriptional complex in the ascidian PCP pathway and its role in axis elongation and cellular intercalation. To this aim, it will be fundamental to identify the other putative AP-1 members responsible together with Fos for notochord correct development.

CONCLUSION

The number of proteins forming the AP-1 complex and their variable combination in specific tissues always strongly impaired the ability to clarify their mechanism of action during normal and neoplastic development. In this study, we evidenced the presence of non-duplicated AP-1 gene members and placed the foundations for understanding the complexity of the AP-1 transcriptional family during embryonic development in ascidians.

We contributed to understanding their potential function in controlling the correct structural organization of cells. Full comprehension of *Ciona* AP-1 transcriptional complex and of its role in mesenchyme differentiation and notochord intercalation could represent an important turning point to understand their direct function in vertebrate erythropoiesis and notochord formation.

Further comparative studies among ascidians and other chordates about AP-1 functional role in PCP movements could reveal interesting similarities as well as novel and significant differences in the mechanisms underlying chordate axis elongation.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://www.aniseed.cnrs.fr/.

AUTHOR CONTRIBUTIONS

PM, FS, and FB did the molecular experiments. LA did the bioinformatics analyses. PM, FS, LA, and AL drafted the manuscript and prepared the figures. AL supervised the experiments and edited and revised the manuscript. MC supervised the bioinformatics analyses and revised the manuscript. All authors approved the manuscript for publication.

ACKNOWLEDGMENTS

We are grateful to Anna Di Gregorio for the *Fos* clone of the Gateway-compatible Unigene collection. Special thanks to Alessandro Amoroso for his help in construct preparation and transgenesis experiments. Our thanks to the RIMAR Department of the Stazione Zoologica Anton Dohrn and, in particular, to the Molecular Biology and Sequencing and the Bioinformatics Units for technical support; to the Confocal Microscopy and Bioimaging Service; and to Alberto Macina, the Marine Biological Resources and the IRM Units for animals fishing and husbandry.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.709696/full#supplementary-material

REFERENCES

Ambrosino, L., Ruggieri, V., Bostan, H., Miralto, M., Vitulo, N., Zouine, M., et al. (2018). Multilevel comparative bioinformatics to investigate evolutionary relationships and specificities in gene annotations: an example for tomato and grapevine. *BMC Bioinformatics* 19:435. doi: 10.1186/s12859-018-2420-y

Amoutzias, G., Veron, A., Weiner, J., Robinson-Rechavi, M., Bornberg-Bauer, E., Oliver, S., et al. (2006). One billion years of bZIP transcription factor evolution:

conservation and change in dimerization and DNA-binding site specificity. *Mol. Biol. Evol.* 24, 827–835. doi: 10.1093/molbev/msl211

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *BBA Rev. Cancer* 1072, 129–157. doi: 10.1016/0304-419X(91)90011-9

Angel, P., and Szabowski, A. (2002). Function of AP-1 target genes in mesenchymal-epithelial cross-talk in skin. *Biochem. Pharmacol.* 64, 949–956. doi: 10.1016/S0006-2952(02)01158-9

Behrens, A., Haigh, J., Mechta-Grigoriou, F., Nagy, A., Yaniv, M., and Wagner, E. F. (2003). Impaired intervertebral disc formation in the absence of Jun. *Development* 130, 103–109. doi: 10.1242/dev.00186

Brown, G. R., Hem, V., Katz, K. S., Ovetsky, M., Wallin, C., Ermolaeva, O., et al. (2015). Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res.* 43, D36–D42. doi: 10.1093/nar/gku1055

Brozovic, M., Dantec, C., Dardaillon, J., Dauga, D., Faure, E., Gineste, M., et al. (2018). ANISEED 2017: Extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. *Nucleic Acids Res.* 46, D718–D725. doi: 10.1093/nar/gkx1108

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics* 10:4291. doi: 10.1186/1471-2105-10-421

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 566, 496–502. doi: 10.1038/s41586-019-0969-x

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/BIOINFORMATICS/BTP348

Christiaen, L., Wagner, E., Shi, W., and Levine, M. (2009). Whole-mount in situ hybridization on sea squirt (Ciona intestinalis) embryos. *Cold Spring Harb. Protoc.* 4:db.rot5348. doi: 10.1101/pdb.prot5348

Ciaccio, P. J., and Tew, K. D. (1994). cDNA and deduced amino acid sequences of a human colon dihydrodiol dehydrogenase. *BBA Bioenerg.* 1186, 129–132. doi: 10.1016/0005-2728(94)90144-9

Corbo, J. C., Levine, M., and Zeller, R. W. (1997). Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. *Development* 124, 589–602. doi: 10.1242/dev.124.3.589

D'Aniello, E., Pezzotti, M. R., Locascio, A., and Branno, M. (2011). Onecut is a direct neural-specific transcriptional activator of Rx in Ciona intestinalis. *Dev. Biol.* 355, 358–371. doi: 10.1016/j.ydbio.2011.05.584

Eferl, R., Sibilia, M., Hilberg, F., Fuchsbichler, A., Kufferath, I., Guertl, B., et al. (1999). Functions of c-Jun in liver and heart development. *J. Cell Biol.* 145, 1049–1061. doi: 10.1083/jcb.145.5.1049

Goode, D. K., Obier, N., Vijayabaskar, M. S., Lie-A-Ling, M., Lilly, A. J., Hannah, R., et al. (2016). Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev. Cell* 36, 572–587. doi: 10.1016/j.devcel.2016.01.024

Hagberg, A., Schult, D. A., and Swart, P. J. (2008). *Exploring Network Structure, Dynamics, and Function Using NetworkX*. Available online at: http://conference.scipy.org/proceedings/SciPy2008/paper_2 (accessed March 25, 2021).

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., et al. (2007). A web-based interactive developmental table for the Ascidian Ciona intestinalis, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. *Dev. Dyn.* 236, 1790–1805. doi: 10.1002/dvdy.21188

Imai, K. S., Hino, K., Yagi, K., Satoh, N., and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* 131, 4047–4058. doi: 10.1242/dev.01270

Jiang, D., Munro, E. M., and Smith, W. C. (2005). Ascidian prickle regulates both mediolateral and anterior-posterior cell polarity of notochord cells. *Curr. Biol.* 15, 79–85. doi: 10.1016/j.cub.2004.12.041

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

José-Edwards, D. S., Kerner, P., Kugler, J. E., Deng, W., Jiang, D., and Di Gregorio, A. (2011). The identification of transcription factors expressed in the notochord of Ciona intestinalis adds new potential players to the brachyury gene regulatory network. *Dev. Dyn.* 240, 1793–1805. doi: 10.1002/dvdy.22719

José-Edwards, D. S., Oda-Ishii, I., Kugler, J. E., Passamaneck, Y. J., Katikala, L., Nibu, Y., et al. (2015). Brachyury, Foxa2 and the cis-regulatory origins of the notochord. *PLoS Genet.* 11:e1005730. doi: 10.1371/journal.pgen.1005730

Karin, M., Liu, Z. G., and Zandi, E. (1997). AP-1 function and regulation. *Curr. Opin. Cell Biol.* 9, 240–246. doi: 10.1016/S0955-0674(97)80068-3

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

Kenner, L., Hoebertz, A., Beil, T., Keon, N., Karreth, F., Eferl, R., et al. (2004). Mice lacking JunB are osteopenic due to cell-autonomous osteoblast and osteoclast defects. *J. Cell Biol.* 164, 613–623. doi: 10.1083/jcb.200308155

Keys, D. N., Levine, M., Harland, R. M., and Wallingford, J. B. (2002). Control of intercalation is cell-autonomous in the notochord of Ciona intestinalis. *Dev. Biol.* 246, 329–340. doi: 10.1006/dbio.2002.0656

Klemm, J. D., Schreiber, S. L., and Crabtree, G. R. (1998). Dimerization as a regulatory mechanism in signal transduction. *Annu. Rev. Immunol.* 16, 569–592. doi: 10.1146/annurev.immunol.16.1.569

Kobayashi, K., Maeda, K., Tokuoka, M., Mochizuki, A., and Satou, Y. (2018). Controlling cell fate specification system by key genes determined from network structure. *Science* 4, 281–293. doi: 10.1016/j.isci.2018.05.004

Lee, C. R., Sakai, D., Nakai, T., Toyama, K., Mochida, J., Alini, M., et al. (2007). A phenotypic comparison of intervertebral disc and articular cartilage cells in the rat. *Eur. Spine J.* 16, 2174–2185. doi: 10.1007/s00586-007-0475-y

Lee, S. Y., Yoon, J., Lee, M. H., Jung, S. K., Kim, D. J., Bode, A. M., et al. (2012). The role of heterodimeric AP-1 protein comprised of JunD and c-Fos proteins in hematopoiesis. *J. Biol. Chem.* 287, 31342–31348. doi: 10.1074/jbc.M112.387266

Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature* 424, 147–151. doi: 10.1038/nature01763

Li, P., Lahvic, J. L., Binder, V., Pugach, E. K., Riley, E. B., Tamplin, O. J., et al. (2015). Epoxyeicosatrienoic acids enhance embryonic haematopoiesis and adult marrow engraftment. *Nature* 523, 468–471. doi: 10.1038/nature14569

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S. A., et al. (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* 31, 4318–4333. doi: 10.1038/emboj.2012.275

Liebermann, D. A., Gregory, B., and Huffman, B. (1998). AP-1 (Fos/Jun) transcription factors in hematopoietic differentiation and apoptosis (Review). *Int. J. Oncol.* 12, 685–700. doi: 10.3892/ijo.12.3.685

Locascio, A., Aniello, F., Amoroso, A., Manzanares, M., Krumlauf, R., and Branno, M. (1999). Patterning the ascidian nervous system: structure, expression and transgenic analysis of the CiHox3 gene. *Development* 126, 4737–4748. doi: 10.1242/dev.126.21.4737

Macera, M. J., Szabo, P., Wadgaonkar, R., Siddiqui, M. A. Q., and Verma, R. S. (1992). Localization of the gene coding for ventricular myosin regulatory light chain (MYL2) to human chromosome 12q23-q24.3. *Genomics* 13, 829–831. doi: 10.1016/0888-7543(92)90161-K

Niwano, T., Takatori, N., Kumano, G., and Nishida, H. (2009). Wnt5 is required for notochord cell intercalation in the ascidian Halocynthia roretzi. *Biol. Cell* 101, 645–659. doi: 10.1042/bc20090042

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189

Obier, N., Cauchy, P., Assi, S. A., Gilmour, J., Lie-A-Ling, M., Lichtinger, M., et al. (2016). STEM CELLS AND REGENERATION Cooperative binding of AP-1 and TEAD4 modulates the balance between vascular smooth muscle and hemogenic cell fate. *Development.* 143, 4324–4340. doi: 10.1242/dev.139857

Pereira, C. F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C. E., et al. (2013). Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* 13, 205–218. doi: 10.1016/j.stem.2013.05.024

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/JOURNAL.PONE.0009490

Ranea, J. A. G., Grant, A., Thornton, J. M., and Orengo, C. A. (2005). Microeconomic principles explain an optimal genome size in bacteria. *Trends Genet.* 21, 21–25. doi: 10.1016/j.tig.2004.11.014

Sakai, L. Y., Keene, D. R., and Engvall, E. (1986). Fibrillin, a new 350-kD glycoprotein, is a component of extracellular microfibrils. *J. Cell Biol.* 103, 2499–2509. doi: 10.1083/jcb.103.6.2499

Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., et al. (2002). A cDNA resource from the basal chordateCiona intestinalis. *Genesis* 33, 153–154. doi: 10.1002/gene.10119

Schwab, J. H., Boland, P. J., Agaram, N. P., Socci, N. D., Guo, T., O'Toole, G. C., et al. (2009). Chordoma and chondrosarcoma gene profile: implications for immunotherapy. *Cancer Immunol. Immunother.* 58, 339–349. doi: 10.1007/s00262-008-0557-7

Segade, F., Cota, C., Famiglietti, A., Cha, A., and Davidson, B. (2016). Fibronectin contributes to notochord intercalation in the invertebrate chordate, *Ciona intestinalis*. *Evodevo* 7:21. doi: 10.1186/s13227-016-0056-4

Sherwood, D. R., Butler, J. A., Kramer, J. M., and Sternberg, P. W. (2005). FOS-1 promotes basement-membrane removal during anchor-cell invasion in C. elegans. *Cell* 121, 951–962. doi: 10.1016/j.cell.2005.03.031

Shukunami, C., and Hiraki, Y. (2001). Role of cartilage-derived anti-angiogenic factor, chondromodulin-I, during endochondral bone formation. *Osteoarthr. Cartil* 9, S91–S101. doi: 10.1053/joca.2001.0450

Spagnuolo, A., and Di Lauro, R. (2002). Cititf1 and endoderm differentiation in Ciona intestinalis. *Gene* 287, 115–119. doi: 10.1016/S0378-1119(01)00830-7

Szalóki, N., Krieger, J. W., Komáromi, I., Tóth, K., and Vámosi, G. (2015). Evidence for homodimerization of the c-fos transcription factor in live cells revealed by fluorescence microscopy and computer modeling. *Mol. Cell. Biol.* 35, 3785–3798. doi: 10.1128/mcb.00346-15

Tokuoka, M., Satoh, N., and Satou, Y. (2005). A bHLH transcription factor gene, Twist-like1, is essential for the formation of mesodermal tissues of Ciona juveniles. *Dev. Biol.* 288, 387–396. doi: 10.1016/j.ydbio.2005.09.018

Topczewski, J., Sepich, D. S., Myers, D. C., Walker, C., Amores, A., Lele, Z., et al. (2001). The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension. *Dev. Cell* 1, 251–264. doi: 10.1016/S1534-5807(01)00005-3

Van Nimwegen, E. (2003). Scaling laws in the functional content of genomes. *Trends Genet.* 19, 479–484. doi: 10.1016/S0168-9525(03)00203-8

Wagner, E. F. (2001). Ap-1 – introductory remarks. *Oncogene* 20, 2334–2335. doi: 10.1038/sj.onc.1204416

Wallingford, J. B., Rowning, B. A., Vogell, K. M., Rothbächer, U., Fraser, S. E., and Harland, R. M. (2000). Dishevelled controls cell polarity during Xenopus gastrulation. *Nature* 405, 81–85. doi: 10.1038/35011077

Wang, Z. Q., Ovitt, C., Grigoriadis, A. E., Möhle-Steinlein, U., Rüther, U., and Wagner, E. F. (1992). Bone and haematopoietic defects in mice lacking c-fos. *Nature* 360, 741–745. doi: 10.1038/360741a0

Wei, Y., Ma, D., Gao, Y., Zhang, C., Wang, L., and Liu, F. (2014). Ncor2 is required for hematopoietic stem cell emergence by inhibiting Fos signaling in zebrafish. *Blood* 124, 1578–1585. doi: 10.1182/blood-2013-11-541391

Xia, Y., Brown, L., Yang, C. Y. C., Tsan, J. T., Siciliano, M. J., Espinosa, R., et al. (1991). TAL2, a helix-loop-helix gene activated by the (7;9)(q34;q32) translocation in human T-cell leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 88, 11416–11420. doi: 10.1073/pnas.88.24.11416

Yamada, L., Kobayashi, K., Degnan, B., Satoh, N., and Satou, Y. (2003). A genomewide survey of developmentally relevant genes in Ciona intestinalis. IV. Genes for HMG transcriptional regulators, bZip and GATA/Gli/Zic/Snail. *Dev. Genes Evol.* 213, 245–253. doi: 10.1007/s00427-003-0316-x

Zeitlinger, J., Kockel, L., Peverali, F. A., Jackson, D. B., Mlodzik, M., and Bohmann, D. (1997). Defective dorsal closure and loss of epidermal decapentaplegic expression in Drosophila fos mutants. *EMBO J.* 16, 7393–7401. doi: 10.1093/emboj/16.24.7393

Zeller, R. W., Weldon, D. S., Pellatiro, M. A., and Cone, A. C. (2006). Optimized green fluorescent protein variants provide improved single cell resolution of transgene expression in ascidian embryos. *Dev. Dyn.* 235, 456–467. doi: 10.1002/dvdy.20644

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Marotta, Salatiello, Ambrosino, Berruto, Chiusano and Locascio. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 12 August 2021
doi: 10.3389/fcell.2021.702290

A Potential Method for Rapid Screening of Amphioxus Founder Harboring Germline Mutation and Transgene

Jiaqi Zou†*, Xiaotong Wu*†*, Chenggang Shi, Yanhong Zhong, Lei Zhang, Qiuning Yan, Liuru Su and Guang Li**

State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen, China

OPEN ACCESS

Edited by:
Salvatore D'Aniello,
Zoological Station Anton Dohrn, Italy

Reviewed by:
Simona Candiani,
University of Genoa, Italy
Vladimir Soukup,
Charles University, Czechia

Correspondence:
Guang Li
guangli@xmu.edu.cn

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *29 April 2021*
Accepted: *22 July 2021*
Published: *12 August 2021*

Citation:
Zou J, Wu X, Shi C, Zhong Y, Zhang L, Yan Q, Su L and Li G (2021) A Potential Method for Rapid Screening of Amphioxus Founder Harboring Germline Mutation and Transgene. Front. Cell Dev. Biol. 9:702290. doi: 10.3389/fcell.2021.702290

Amphioxus is a promising model organism for understanding the origin and evolution of vertebrates due to its basal phylogenetic position among chordates. We here compared the mutation efficacy and mutation type of tail tips and gametes of amphioxus founders injected with Cas9 protein and six different sgRNAs targeting five distinct genes, and revealed a strong correlation for mutation efficacy and a mild correlation for mutation type among the two tissues. In addition, we also observed a positive relationship between gene insertions observed in tail tips and gametes of amphioxus founders injected with Tol2 transposase and two different transgenic constructs. Finally, we showed that amphioxus larvae which had their tail tips cut at the 3–4 gill-slit stage were able to recover within 6 days and developed a normal number of gonads at the adult stage, and that F0 larvae carry similar mutation efficacy and type in the posterior end to that in the tail tips after their metamorphosis. Together, these findings suggest a great potential for obtaining valid amphioxus founders with desired mutations and transgenes at as early as the early larval stage, which will certainly speed up the generation of amphioxus mutants and transgenes and make it more cost- and labor-effective.

Keywords: amphioxus, mutation, transgene, tail tissue, gametes, correlation

INTRODUCTION

The phylum Chordata constitutes of three subphyla: Cephalochordata, Urochordata and Vertebrata. Among them, Cephalochordata is the most basal subphylum (Delsuc et al., 2006) which comprises around thirty living species, commonly known as lancelets or amphioxus. Anatomically, amphioxus have segmented muscles, a dorsal hollow neural tube, notochord, gill slits, a through gut and postanal tail while lacking bones, appendages, paired sensory organs (such as eyes and ears), an adaptive immune system, and a well-developed brain and visceral organs. Developmentally, amphioxus embryos are similar to vertebrate embryos in the sense that they undergo blastulation, gastrulation and neurulation. Nevertheless, they are like a single-cell layered ball at the blastula stage, gastrulate simply by invagination of the vegetal pole toward the animal pole, and do not form neural crest cells and definitive placode at the neurula stage. Genetically, the amphioxus genome has retained many ancestral chordate traits but has not undergone massive duplications like that in vertebrates (Putnam et al., 2008). Due to the above-mentioned advantages, amphioxus has been considered as a promising model organism for studying the origin of vertebrates and the evolution of chordates since they were discovered in the 1,800s (Bertrand and Escriva, 2011).

The thirty described amphioxus species are divided into three genera: *Branchiostoma*, *Asymmetron,* and *Epigonichthys*, and most of them belong to genus *Branchiostoma* (Poss and Boschung, 1996). Up to date, the genomes of three species of *Branchiostoma*, *B. floridae*, *B. belcheri,* and *B. lanceolatum*, have been sequenced (Putnam et al., 2008; Huang et al., 2014; Marletaz et al., 2018). In addition, methods for year-round spawning induction has been developed for all three species (Benito-Gutierrez et al., 2013; Li et al., 2014b, 2017), and TALEN-based genome editing and Tol2 transposase-mediated transgenic methods have been introduced in *B. floridae* and *B. belcheri* (Li et al., 2014a, 2017; Shi et al., 2018). Very recently, the CRIPSR/Cas9 system has also been applied for genome editing in *B. floridae* (Su et al., 2020). These advances have greatly accelerated the rate of establishing amphioxus as a model organism, and enables researchers to robustly dissect the function of amphioxus genes (Hu et al., 2017; Li et al., 2017; Ren et al., 2020; Zhong et al., 2020; Zhu et al., 2020). However, current methods used for identifying valid founder amphioxus (F0) carrying desired mutations or transgene are time-consuming (Holland and Li, 2021) since they rely on animal spawning that normally takes at least 3 months (from their births) for *B. floridae* or more for other amphioxus species (Zhang et al., 2007; Escriva, 2018). Moreover, to ensure success, especially for transgene and mutation sites of low efficacy, more than several dozen F0 animals need to be raised. This would take more effort and slow down the growth rate of the founders due to high culture density. To overcome this drawback, we here compared the mutation efficacy and mutation type induced by Cas9/sgRNA ribonucleoprotein (RNP) among tail tips and gametes of F0 individuals, and found that both of the two parameters are well correlated between the two tissues. We also observed a positive relationship of successful transgene insertion between tail tips and gametes of F0 individuals injected with Tol2 transposase and two vectors. Finally, we demonstrate that amphioxus larvae whose tail tips were cut at 3 gill-slit stage were able to recover within 6 days and develop a normal number of gonads at the adult stage, and that F0 larvae carry similar mutation efficacy and type in the posterior end to that in the tail tips after their metamorphosis. These results together indicate that it is possible to identify F0 amphioxus carrying desired mutations or transgenes at as early as the 3 gill-slit stage. This advantage will greatly accelerate the generation of genetically modified amphioxus animals.

MATERIALS AND METHODS

Animal and Embryo Cultivation

Amphioxus *Branchiostoma floridae* were obtained from Dr. Jr-Kai Yu's laboratory at Institute of Cellular and Organismic Biology, Academia Sinica, Taiwan. They were maintained and induced to spawn following the protocols as we described before for *B. belcheri* animals (Li et al., 2012, 2014b). Fertilization and embryo cultivation were carried out as previously reported unless otherwise stated (Liu et al., 2013).

Cultivation of Amputated Larvae

The posterior end of amphioxus larvae (with 3–4 gill slits) were amputated with a double-edged blade under a stereoscope. They were then raised in a petri dish (diameter = 6 cm) containing filtered sea water placed in a 30°C incubator and fed with *Dicrateria zhangjiangensis* twice a day. Sea water was changed once per day. After metamorphosis, they were moved to a 5 L plastic barrel and raised as previously described (Li et al., 2012, 2014b). Photographs were taken with an inverted microscope (Olympus, IX71) or an SZX10 fluorescent stereoscope (Olympus, Japan).

Mutant and Transgenic Founder Generation

We used the CRISPR/Cas9 or Tol2 system reported previously to generate mutant or transgenic founders (Shi et al., 2018; Su et al., 2020). Seven sgRNAs, which target *Invs* (one sgRNA), *Wnt3* (one sgRNA), *VegT* (one sgRNA), *Mop* (one sgRNA), *Cyp19a2* (two sgRNAs), and *Tesd* (one sgRNA) genes, respectively, and two *mCherry* transgenic constructs, that, respectively, include the promoter (1,784 bp) of the *Xenopus laevis Slug* gene (Vallin et al., 2001) and that (3,285 bp) of the *B. floridae Pou4* gene, were used. gRNA targeting sequences and primers used for transgenic vector construction were, respectively, listed in **Supplementary Tables 1, 2**.

Mutation Efficacy and Mutation Type Detection in Tail Tips and Gametes of F0 Amphioxus

F0 founders were crossed with wild type amphioxus to generate F1 embryos. For mutation efficacy detection, the tail tip of F0 amphioxus, semen from male mutants and F1 gastrula embryos (around 50) of female founders were lysed to get genomic DNA using the Animal Tissue Direct PCR kit (FOREGENE, Chengdu, China). PCR amplification and restriction enzyme digestion assay was performed as previously described (Li et al., 2017). Mutation efficacy was estimated by comparing band intensity between uncut and all bands [uncut/(uncut + cut)]. Band intensity was quantified using software implemented in the Tanon Gis system (Tanon, Shanghai, China) (Su et al., 2020). For mutation type detection, the PCR products obtained above were further cloned into the pGEM-T easy vector (Promega, United States) (except *Tesd* gene, whose PCR products were ligated after digestion with *Eae*I). Clones with mutated target site were identified as described above using restriction enzyme digestion assay, and then sequenced to determine their mutation types. Primers and restriction enzymes used for each gRNA were, respectively, shown in **Supplementary Tables 1, 3**.

Genotyping of Transgenic Amphioxus

Genomic DNA of tail tips and gametes of transgenic F0 founders was obtained as described above, and used as the template to amplify a 624 bp fragment of pMini-mCherry backbone with primers mCherry-F1 (5′-CTTCGCCTGGGACATCCTGT-3′) and mCherry-R1 (5′-GCATTCTAGTTGTGGTTTGTC-3′) under the following conditions: 95°C for 5 min, 38 × (95°C for 30 s, 62°C for 20 s, 72°C for 30 s), 72°C for 5 min, and hold at 4°C. Ten microliters of each PCR product was used for gel electrophoresis and those including the 624 bp band indicated successful insertion of the target sequence.

Statistical Analysis

Data were analyzed using a two-tailed unpaired *t*-test, correlation or simple linear regression with GraphPad Prism 8.3.0. We set gamete efficacy as x value and tail efficacy as y value for correlation analysis of mutation efficacy. And for correlation analysis of mutation type, in each individual, we regard the proportion of one gamete mutation type as x value and the same type of tail tip as y value.

RESULTS AND DISCUSSION

The Mutation Efficacy Measured in Tail Tips of F0 Amphioxus Correlates Strongly With That in Their Gametes

Both male and female amphioxus have more than twenty pairs of gonads, which are distributed along the ventral tip of the myoseptal walls in the middle part of the body. Staining of marker genes suggests that amphioxus primordial germ cells

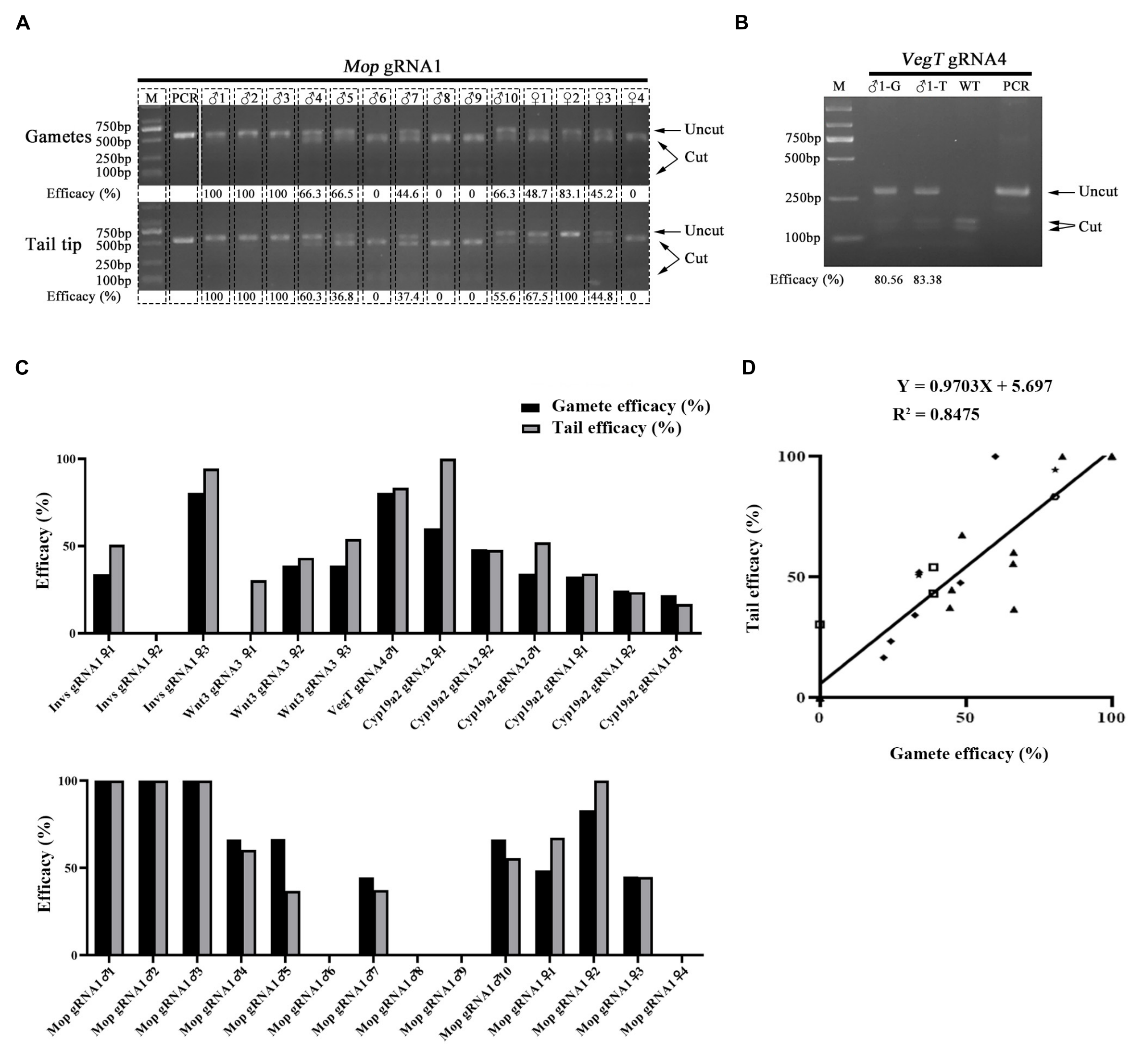


FIGURE 1 | Correlation of mutation efficacy between tail tips and gametes of F0 individuals. **(A,B)** Mutation detection in tail tips (T) and gametes (G) of F0 individuals injected with *Mop* gRNA1 **(A)** and *VegT* gRNA4 **(B)**. The induced mutation efficacies (estimated as percentages of uncut PCR products) are labeled under the gel image. **(C)** Summary of mutation efficacies detected in tail tips and gametes of all F0 individuals examined. **(D)** Correlation analysis of efficacies between tail tips and gametes of F0 individuals. Efficacy of each gRNA is marked with a different symbol. The linear equation is $Y = 0.9703X + 5.697$, $R^2 = 0.8475$.

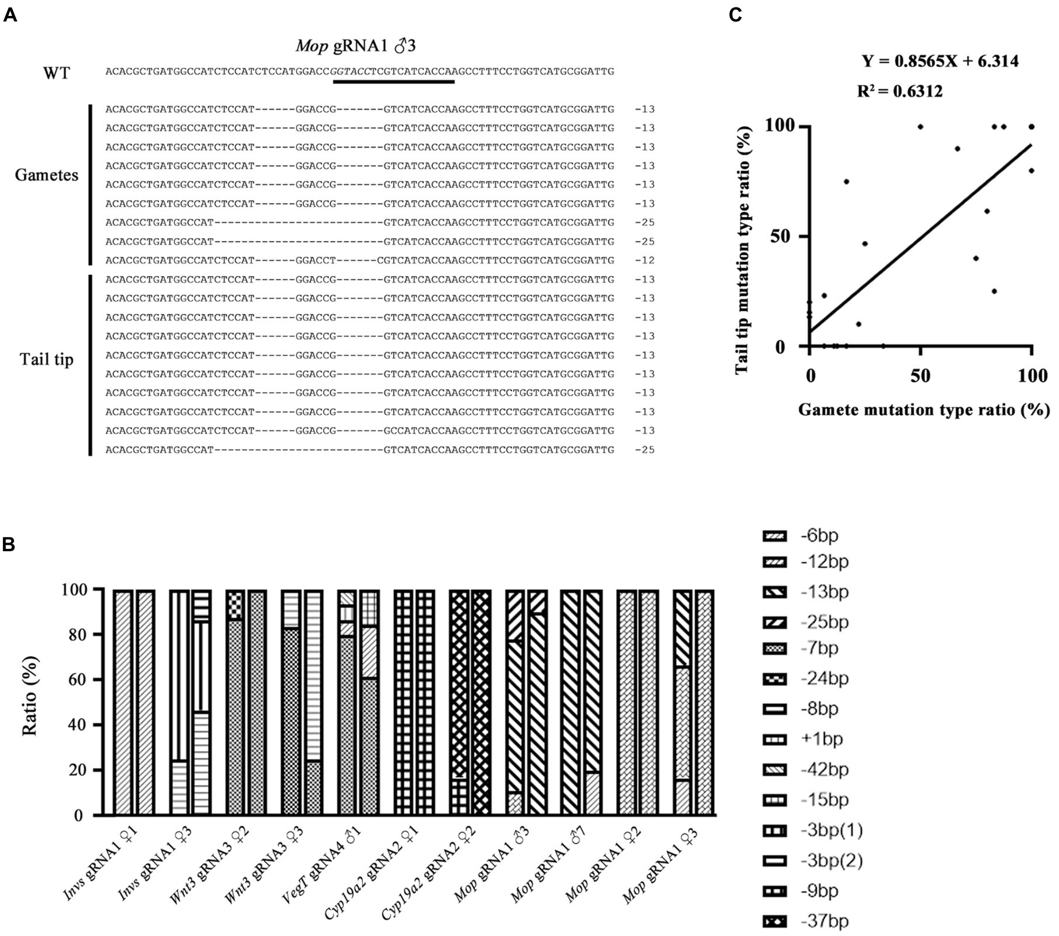


FIGURE 2 | Correlation of mutation types between tail tips and gametes of F0 individuals. **(A)** Mutation types detected in tail tip and gametes of a F0 male (♂3) injected with *Mop* gRNA1. The gRNA target site is underlined and the restriction enzyme recognition sequences is marked in italics. **(B)** Summary of mutation types detected in gametes and tails of all F0 individuals. For each individual, the first column represents mutation types detected in the gametes and the second column represents those detected in the tail tips. Each type of mutation is marked with a different symbol. **(C)** A linear correlation analysis on the percentage of each mutation type between tails and gametes. For those with multiple mutation types, each type was treated as an independent dataset. The linear equation is $Y = 0.8565X + 6.314$, $R^2 = 0.6312$.

(PGCs), from which the gonad cells derive, are located within the tail bud and proliferate together with the tail bud cells from the mid-neurula (8 pairs of somites) stage (Zhang et al., 2013). During posterior somite formation from the tail bud, some of these PGCs are thought to be deposited near the forming myomere boundaries and then settle at the ventral tip of the myoseptal walls during subsequent development (Zhang et al., 2013). This finding shows a close relationship between tail and gonad tissue during amphioxus development, and raises a possibility that gonads of F0 animals injected with RNP would carry similar mutation efficacy compared to tail tissue. To test this, we examined the mutation efficacy of tail and gonad tissue of F0 individuals injected with Cas9 protein and gRNAs targeting *Invs*, *Wnt3*, *VegT*, *Cyp19a2,* and *Mop* genes. This includes six gRNAs in which two targeting the *Cyp19a2* gene and one targeting each of the other four genes. After being injected (gRNAs for each gene were injected separately), the embryos were raised to adulthood and their tail tips and gametes (for females their offspring generated by crossing them with wild type males were used) were then collected for mutation efficacy detection using a restriction endonuclease assay. In total, three females for *Invs* gRNA1 and *Wnt3* gRNA3, one male for *VegT* gRNA4, two females and one male for *Cyp19a2* gRNA1 and 2, and four females and ten males for *Mop* gRNA1 were analyzed. Among these 24 individuals, five (*Invs* gRNA1 ♀2, *Mop* gRNA1 ♂6, ♂8, ♂9, and ♀4) carried no mutation in either tails or gonads, one showed mutations (30.46% efficacy) in tail tissue but not in gametes, and the remaining 19 exhibited mutations (16.67–100% efficacy) in both tails and gametes (**Figures 1A–C** and **Supplementary Figure 1**). Correlation analysis revealed that the mutation efficacies detected in the gametes of these animals were highly correlated with those detected in their tail tips ($Y = 0.9703X + 5.697$, $R^2 = 0.8475$) (**Figure 1D**).

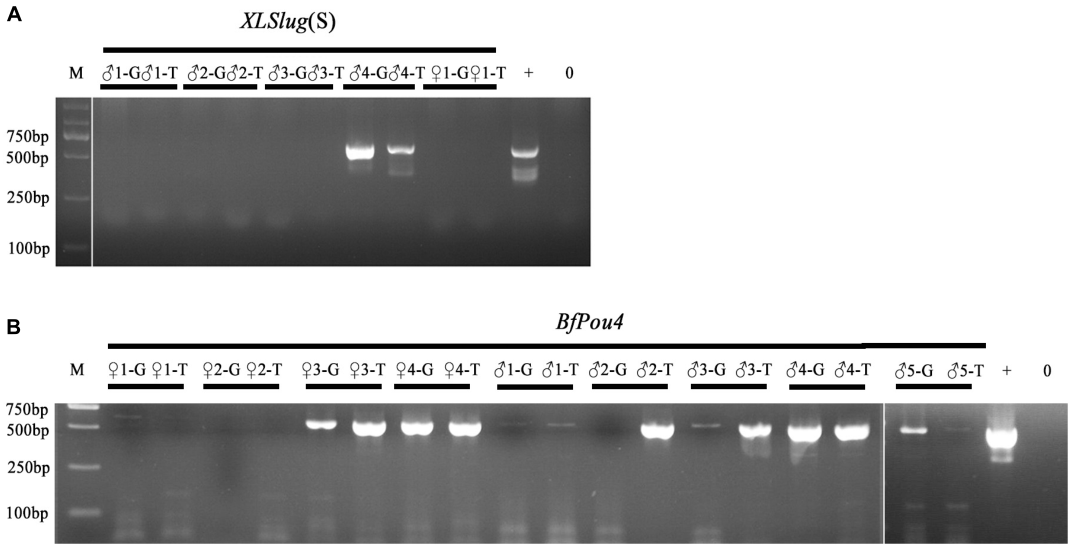


FIGURE 3 | Detection of transgene in tail tips and gametes of F0 animals. **(A)** Analysis of F0 animals injected with *XLSlug*(S)-*mCherry* construct. **(B)** Analysis of F0 animals injected with *BfPou4-mCherry* construct. The presence of the 624 bp band indicates successful insertion of target sequence. + means positive control and − means negative control.

This result indicates that the germline mutation efficacy of F0 amphioxus can be well indicated by the mutation efficacy of their tail tips.

Correlation of Mutation Type Between Amphioxus Tail Tips and Gametes

We then compared the mutation types between the gametes and tail tips of F0 amphioxus to see if they are also correlated. Eleven individuals used in the analysis above were selected. **Figure 2A** showed an example of the mutation types identified in a male injected with *Mop* gRNA1 (*Mop* gRNA1 ♂3). There were, respectively, three and two types of mutations in the gametes and tail tip. Remarkably, two of these mutations (−13 and −25 bp) were shared by both the gametes and tail tip, and the 13 bp deletion was a major type of mutation in both tissues (6/9 = 66.6% in gametes and 9/10 = 90.0% in tail). The −12 bp mutation found in the gametes was not detected in the tail tip, which might be caused by limited number of clones we analyzed. Similar results were also observed in the other ten animals (**Figure 2B** and **Supplementary Figure 2**). In some extreme cases (*Invs* gRNA1 ♀1, *Mop* gRNA1 ♀2, and *Cyp19a2* gRNA1 ♀1), an identical mutation type was detected in both gametes and tail of the animals. We then conducted a linear correlation analysis on the percentage of each mutation type between the two tissues of the animals examined (for those with multiple mutation types, each type was treated as an independent dataset). The result showed that the mutation types present in the gametes and tails were well correlated and unified into a linear equation ($Y = 0.8565X + 6.314$, $R^2 = 0.6312$) (**Figure 2C**). From these, we infer that the mutation types, at least the major one, in the gametes of F0 amphioxus could be predicted by the mutation types present in their tail tips.

Prediction of Germline Transmission of Transgene in F0 Amphioxus by Analyzing Their Tail Tips

The Tol2 transposase-mediated transgenic method has been recently introduced into *B. floridae* (Shi et al., 2018). Screening of F0 founders carrying the desired transgene is currently conducted after the founder animal spawns, by detecting the presence of transgene in their gametes or offspring embryos using PCR method (Holland and Li, 2021). To see if germline transmission of transgene in F0 amphioxus can be predicted by the presence of the transgene in their tail tissue, we conducted PCR analysis on both tissues of 14 individuals injected with *Tol2* mRNA and *XlSlug*(S)-*mCherry* (5 animals) or *BfPou4-mCherry* (9 animals) constructs. The result showed that amongst these animals, seven (*XlSlug*(S) ♂4, *BfPou4* ♀3, ♀4, ♂1, ♂3, ♂4, and ♂5) contained the transgene in both gametes and tail tips, one (*BfPou4* ♂2) in only the tail but not gametes, while we failed to detect the transgene in the remaining six [*XlSlug*(S) ♂1, ♂2, ♂3, ♀1, and *BfPou4* ♀1 and ♀2] in either gametes or tail tissue (**Figure 3**). Since we could not directly determine transgenic efficacy in the tail tip, we did not know if it is well correlated with that in the gametes. However, the above analysis suggests a strong correlation in the presence or absence of a transgene between the gametes and tail tips of F0 amphioxus. This demonstrates that the germline transmission of transgene in F0 amphioxus could be reliably predicted by detecting the transgene in their tail tips.

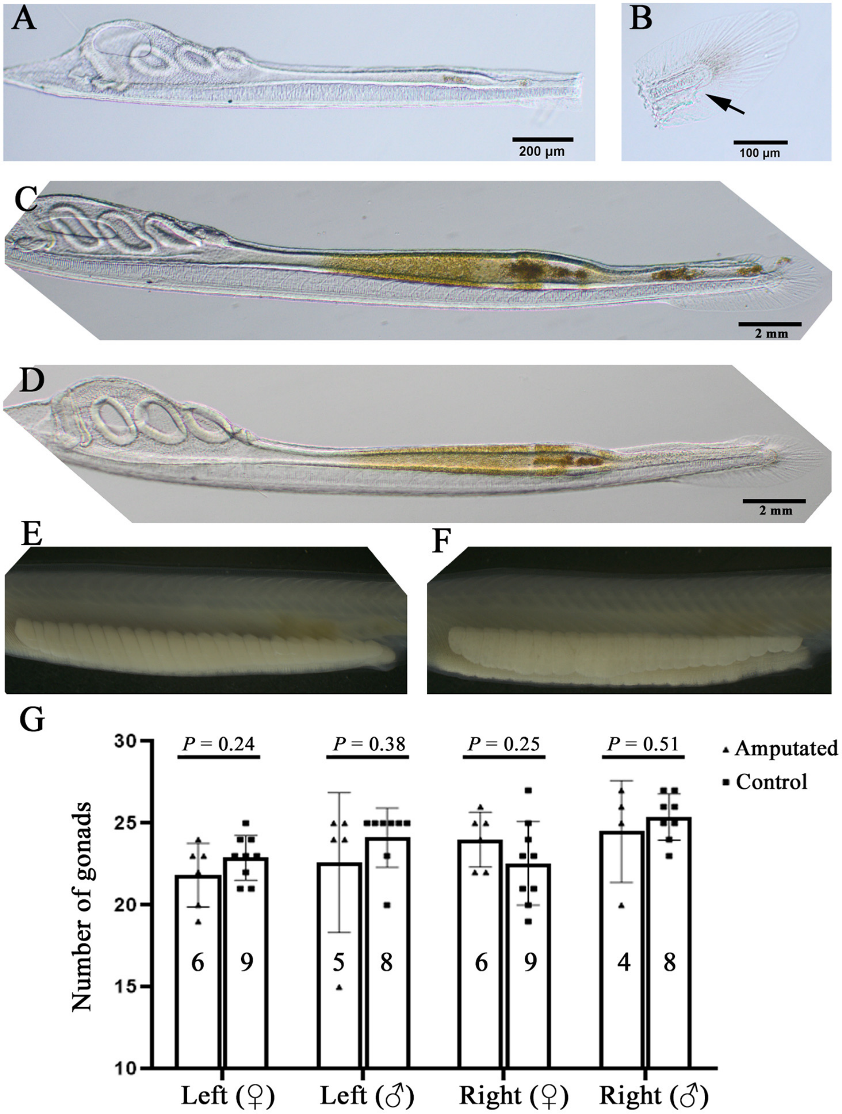


FIGURE 4 | Amphioxus larvae that have undergone posterior end amputation are able to regenerate and develop a normal number of gonads. **(A)** A representative larva just after posterior end amputation. **(B)** An example of the amputated posterior end. Arrow marks the anus. **(C)** An example of larva 3 days after posterior end amputation. Note that its posterior end was almost regenerated. **(D)** A larva with fully recovered posterior end 6 days after amputation. Left-side gonads of a male **(E)** and a female **(F)** whose posterior end was amputated at the larval stage. **(G)** Statistical analysis of gonad number in amputated and wild type (control) animals. Males and females, and their left-side and right-side gonads are compared separately. Animal numbers examined are shown in the columns. *P*-values calculated using a two-tailed unpaired *t*-test are shown above the columns.

Amphioxus Larvae With the Posterior End Removed Are Able to Survive to Adulthood and Grow a Normal Number of Gonads

The results shown above demonstrate that F0 amphioxus have a similar mutation or transgene insertion rate between their gametes and tail tissue, echoing the close relationship between PGCs and tail bud in early development. This finding indicates a possibility to identify F0 amphioxus carrying valid germline mutations and transgenes at early stages before sexual maturity by assaying gene editing efficiency using tail tissue. As an initial step to test the above possibility, we amputated the posterior ends of 14 amphioxus larvae with 3–4 gill slits which include the tail bud and most of the tail fins (slightly anterior to the anus) (**Figures 4A,B** and **Supplementary Figure 3A**) which contained enough cells for genotyping. We found that most of these larvae could recover partially 3 days after amputation, and regenerate the entire posterior end (including the tail bud, anus and tail fin) 6 days after amputation. This demonstrates that amphioxus larvae have strong posterior end regeneration potential and are able to recover after amputation at a much faster rate than adults (Somorjai et al., 2012; **Figures 4C,D** and **Supplementary Figures 3B,C**). Eleven of the 14 amputated

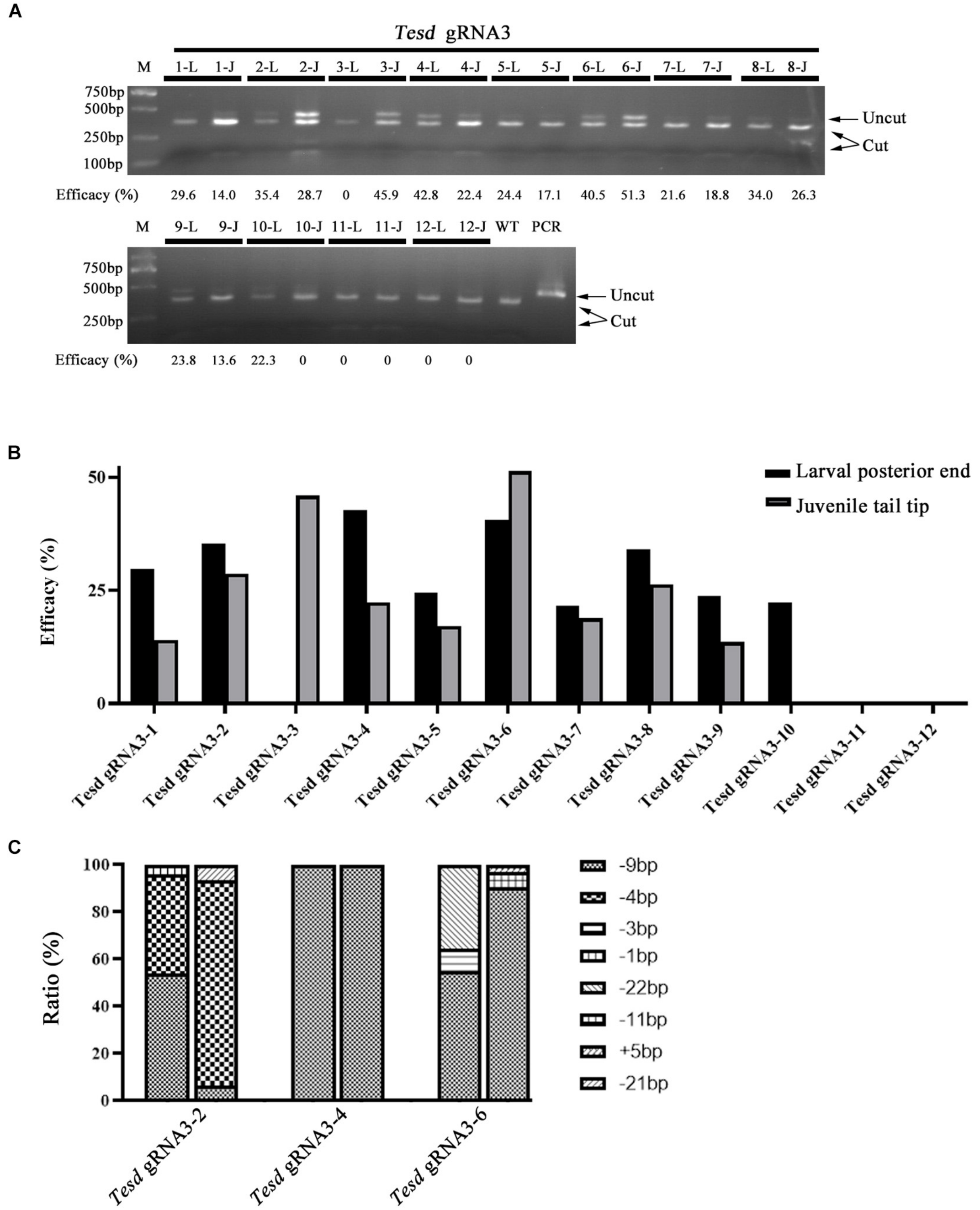


FIGURE 5 | Mutation efficacy and type detected in posterior ends of 4-gill-slits larvae and their tail tips after metamorphosis. **(A)** Mutation efficacy detection by restriction enzyme assay. Twelve F0 individuals injected with Cas9 protein and *Tesd* gRNA3 were examined. The induced mutation efficacies (estimated as percentages of uncut PCR products) are labeled under the gel image. L represents efficacy detected in the posterior end of each individual at the larval stage, and J represents efficacy detected in their tail tips at the juvenile stage. **(B)** Graphic view of the mutation efficacies shown in **A**. **(C)** Mutation types detected in the posterior ends and tail tips of three F0 individuals (#2, #4, and #6). For each individual, the first column represents mutation types detected in the posterior ends and the second column represents those detected in the tail tips. Each type of mutation is marked with a different symbol.

larvae survived to adulthood. Notably, they all grew a normal number of gonads on either side, similar to the untreated wild type animals (**Figures 4E–G** and **Supplementary Figure 3D**). This is unexpected, because the larvae used in this analysis (with 3–4 gill slits) possess only around 20 pairs of somites (Schubert et al., 2001), and amphioxus PGCs are deposited near the forming myomere boundaries only after the mid-neurula stage (with 8 pairs of somite) (Zhang et al., 2013). Moreover, the amputated posterior end includes the whole tail bud and the PGCs within it. We expected the amputated animals to develop around 20-8 = 12 pairs of gonads under the 8th–20th myomeres. To explain this paradox, we suggest that the amputated larvae may have also regenerated PGCs during recovery of the posterior end.

The Posterior End of 4-Gill-Slits F0 Larvae Carries Similar Mutation Efficacy and Type to That in Their Tail Tips After Metamorphosis

We further examined if F0 individuals injected with Cas9 protein and gRNAs carry similar mutation efficacies and types in the posterior end at 4-gill-slits stage and tail tips at early juvenile stage. Twelve individuals injected with the *TesD* gRNA were used in the analysis. Among them, ten showed similar mutation efficacies among the two different tissues ($R^2 = 0.6871$), while the other two carried mutations only in the posterior end at 4-gill-slits stage (#10) or only in the tail tip at early juvenile stage (#3) (**Figures 5A,B**). Further genotyping analysis of 3 individuals (#2, #4, and #6) revealed that they all carried similar major mutation types between the posterior ends at 4-gill-slits stage and tail tips at early juvenile stage ($R^2 = 0.5834$) (**Figure 5C** and **Supplementary Figure 3A**).

CONCLUSION

In summary, we here show that the mutation efficacy and type, and the presence (or absence) of transgene in F0 amphioxus gametes are strongly correlated with and can be reliably predicted by that in their tail tips. This is probably due to the special PGCs distribution mode in this group of animals. To our knowledge, such correlations have not been reported in other animals, including regularly used model organisms such as zebrafish and mouse. Our result also demonstrates that amphioxus larvae with 3–4 gill slits have strong posterior end regeneration potential and are able to fully recover after amputation within 6 days. Moreover, the amputated larvae can survive to adulthood at very high ratios and develop a normal number of gonads. We also show that F0 animals carry similar mutation efficacies and types in their posterior end at 4-gill-slits stage and their tail tips at early juvenile stage. These results, together with the method recently developed for genotyping amphioxus embryos and larvae with tiny piece of tissues (Holland and Li, 2021), suggest a great potential for obtaining F0 animals with desired mutations and transgenes at as early as the 3 gill-slit larval stage. We believe this will greatly accelerate, simplify and reduce the cost of amphioxus transgenic line establishment.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

GL designed the study and conceived the experiments. JZ, XW, and GL performed the experiments, analyzed the data, compiled the figures, wrote, and revised the manuscript. XW and GL generated Invs and Cyp19a2 mutants. YZ generated Wnt3 and VegT mutants. LS and GL generated Mop mutants. LZ and QY generated XlSlug(S)-mCherry and BfPou4-mCherry transgenic lines, respectively. CS and GL performed larvae regeneration experiment. All authors have read and approved the final version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31872186 and 32070458), the National Natural Science Foundation of China (NSFC)/Research Grants Council (RGC) Joint Research Scheme (No. 32061160471), and the Science Foundation of Xiamen for Young Scientist (No. 2019J01022).

ACKNOWLEDGMENTS

We thank Dr. Yichen Dai for her critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.702290/full#supplementary-material

Supplementary Figure 1 | Correlation between mutation efficacy in tail tips and gametes of F0 individuals. **(A–C)** Mutation detection in tail tips (T) and gametes (G) of F0 individuals injected with *Wnt3* gRNA3 **(A)**, *Cyp19a2* gRNA2 **(B)** and *Cyp19a2* gRNA1 **(C)**. The induced mutation efficacies (estimated as percentages of uncut PCR products) are labeled under the gel image.

Supplementary Figure 2 | Correlation between mutation type in tail tips and gametes of F0 individuals. **(A–J)** Mutation types detected in tail tips and gametes of two F0 females (♀1,♀3) injected with *Invs* gRNA1 **(A,B)**, two F0 females (♀2,♀3) injected with *Wnt3* gRNA3 **(C,D)**, one F0 male (♂1) injected with *VegT* gRNA4 **(E)**, two F0 females (♀1,♀2) injected with *Cyp19a2* gRNA2 (F-G) and three F0 (♂7, ♀2, ♀3) injected with *Mop* gRNA1 **(H–J)**. The gRNA target site is underlined and the restriction enzyme recognition sequences is marked in italics.

Supplementary Figure 3 | Regeneration of amphioxus larvae that have undergone posterior end amputation and their gonad development. Fourteen larvae with 3–4 gill slits were amputated and 11 of them survived to adulthood. **(A)** Nine of the amputated larvae (1–9) and three of the posterior ends (10–11) amputated. The scale bar (200 μm) on 1 also applies to panels 2–9, and that on 10 (100 μm) applies to panel 11 too. **(B)** Thirteen of the amputated larvae 3 days after amputation. The scale bar (2 mm) on 1 applies to all panels. **(C)** All 14 amputated larvae 6 days after amputation. The scale bar (2 mm) on 1 applies to all panels. **(D)** Left- and right-side gonads of the 11 survived animals. Right-side gonads of animal 11 are not shown, but they can be scored from the picture taken from its left side.

Supplementary Figure 4 | Mutation types detected in posterior ends of three 4-gill-slit larvae injected with *Tesd* gRNA3 and tail tips of them after metamorphosis. The gRNA target site is underlined and the restriction enzyme recognition sequences is marked in italics. Numbers after * represents colony numbers carrying each mutation type.

REFERENCES

Benito-Gutierrez, E., Weber, H., Bryant, D. V., and Arendt, D. (2013). Methods for generating year-round access to amphioxus in the laboratory. *PloS One* 8:e71599 doi: 10.1371/journal.pone.0071599

Bertrand, S., and Escriva, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. doi: 10.1038/nature04336

Escriva, H. (2018). My favorite animal, amphioxus: unparalleled for studying early vertebrate evolution. *Bioessays* 40:1800130. doi: 10.1002/bies.201800130

Holland, L. Z., and Li, G. (2021). Laboratory culture and mutagenesis of amphioxus (*Branchiostoma floridae*). *Methods Mol. Biol.* 2219, 1–29. doi: 10.1007/978-1-0716-0974-3_1

Hu, G., Li, G., Wang, H., and Wang, Y. (2017). Hedgehog participates in the establishment of left-right asymmetry during amphioxus development by controlling Cerberus expression. *Development* 144, 4694–4703. doi: 10.1242/dev.157172

Huang, S., Chen, Z., Yan, X., Yu, T., Huang, G., Yan, Q., et al. (2014). Decelerated genome evolution in modern vertebrates revealed by analysis of multiple lancelet genomes. *Nat. Commun.* 5:5896 doi: 10.1038/ncomms6896

Li, G., Feng, J., Lei, Y., Wang, J., Wang, H., Shang, L.-K., et al. (2014a). Mutagenesis at specific genomic loci of amphioxus *Branchiostoma* belcheri Using TALEN Method. *J. Genet. Genom.* 41, 215–219. doi: 10.1016/j.jgg.2014.02.003

Li, G., Liu, X., Xing, C., Zhang, H., Shimeld, S. M., and Wang, Y. (2017). Cerberus-Nodal-Lefty-Pitx signaling cascade controls left-right asymmetry in amphioxus. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3684–3689. doi: 10.1073/pnas.1620519114

Li, G., Shu, Z., and Wang, Y. (2014b). Year-round reproduction and induced spawning of Chinese amphioxus, *Branchiostoma* belcheri, in laboratory. *PloS One* 9:e99264. doi: 10.1371/journal.pone.0099264

Li, G., Xi, Y., Shu, Z. H., Chen, X. Y., and Wang, Y. Q. (2012). Consecutive spawnings of Chinese amphioxus, *Branchiostoma* belcheri, in captivity. *PloS One* 7:e50838. doi: 10.1371/journal.pone.0050838

Liu, X., Li, G., Feng, J., Yang, X., and Wang, Y.-Q. (2013). An efficient microinjection method for unfertilized eggs of Asian amphioxus *Branchiostoma* belcheri. *Dev. Genes Evol.* 223, 269–278. doi: 10.1007/s00427-013-0441-0

Marletaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564, 64–70. doi: 10.1038/s41586-018-0734-6

Poss, S. G., and Boschung, H. T. (1996). Lancelets (*Cephalochordata: Branchiostomatidae*): how many species are valid? *Isr. J. Zool.* 42, 13–66.

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–U1063. doi: 10.1038/nature06967

Ren, Q., Zhong, Y., Huang, X., Leung, B., Xing, C., Wang, H., et al. (2020). Step-wise evolution of neural patterning by Hedgehog signalling in chordates. *Nat. Ecol. Evol.* 4, 1247–1255. doi: 10.1038/s41559-020-1248-9

Schubert, M., Holland, L. Z., Stokes, M. D., and Holland, N. D. (2001). Three amphioxus Wnt genes (AmphiWnt3. AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Dev. Biol.* 240, 262–273. doi: 10.1006/dbio.2001.0460

Shi, C., Huang, J., Chen, S., Li, G., and Wang, Y. (2018). Generation of two transgenic amphioxus lines using the Tol2 transposon system. *J. Genet. Genom.* 45, 513–516. doi: 10.1016/j.jgg.2018.06.002

Somorjai, I. M. L., Somorjai, R. L., Garcia-Fernandez, J., and Escriva, H. (2012). Vertebrate-like regeneration in the invertebrate chordate amphioxus. *Proc. Natl. Acad. Sci.U.S.A.*109, 517–522. doi: 10.1073/pnas.1100045109

Su, L., Shi, C., Huang, X., Wang, Y., and Li, G. (2020). Application of CRISPR/Cas9 nuclease in amphioxus genome editing. *Genes* 11:1311. doi: 10.3390/genes11111311

Vallin, J., Thuret, R., Giacomello, E., Faraldo, M. M., and Broders, F. (2001). Cloning and characterization of threexenopus slug promoters reveal direct regulation by Lef/β-catenin signaling. *J. Biol. Chem.* 276, 30350–30358.

Zhang, Q.-J., Luo, Y.-J., Wu, H.-R., Chen, Y.-T., and Yu, J.-K. (2013). Expression of germline markers in three species of amphioxus supports a preformation mechanism of germ cell development in *Cephalochordates*. *Evodevo* 4:17. doi: 10.1186/2041-9139-4-17

Zhang, Q.-J., Sun, Y., Zhong, J., Li, G., Lue, X.-M., and Wang, Y.-Q. (2007). Continuous culture of two lancelets and production of the second filial generations in the laboratory. *J. Exp. Zool. B Mol. Dev. Evol.* 308B, 464–472. doi: 10.1002/jez.b.21172

Zhong, Y., Herrera-Ubeda, C., Garcia-Fernandez, J., Li, G., and Holland, P. W. H. (2020). Mutation of amphioxus Pdx and Cdx demonstrates conserved roles for ParaHox genes in gut, anus and tail patterning. *BMC Biol.* 18:68 doi: 10.1186/s12915-020-00796-2

Zhu, X., Shi, C., Zhong, Y., Liu, X., Yan, Q., Wu, X., et al. (2020). Cilia-driven asymmetric Hedgehog signalling determines the amphioxus left-right axis by controlling Dand5 expression. *Development* 147:dev182469. doi: 10.1242/dev.182469

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zou, Wu, Shi, Zhong, Zhang, Yan, Su 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) and the copyright owner(s) 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 Cell and Developmental Biology

BRIEF RESEARCH REPORT
published: 18 August 2021
doi: 10.3389/fcell.2021.700860

Genetic Mechanism for the Cyclostome Cerebellar Neurons Reveals Early Evolution of the Vertebrate Cerebellum

Fumiaki Sugahara[1,2]*, Juan Pascual-Anaya[2,3,4], Shigehiro Kuraku[5,6], Shigeru Kuratani[2,7] and Yasunori Murakami[8]

[1] Division of Biology, Hyogo College of Medicine, Nishinomiya, Japan, [2] Evolutionary Morphology Laboratory, RIKEN Cluster for Pioneering Research (CPR), Kobe, Japan, [3] Department of Animal Biology, Faculty of Sciences, University of Málaga, Málaga, Spain, [4] Andalusian Centre for Nanomedicine and Biotechnology (BIONAND), Málaga, Spain, [5] Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Japan, [6] Molecular Life History Laboratory, Department of Genomics and Evolutionary Biology, National Institute of Genetics, Mishima, Japan, [7] Laboratory for Evolutionary Morphology, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Japan, [8] Graduate School of Science and Engineering, Ehime University, Matsuyama, Japan

OPEN ACCESS

Edited by:
David Ellard Keith Ferrier,
University of St Andrews,
United Kingdom

Reviewed by:
Sebastian Shimeld,
University of Oxford, United Kingdom
Thomas Butts,
University of Liverpool,
United Kingdom

***Correspondence:**
Fumiaki Sugahara
fu-sugahara@hyo-med.ac.jp

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *27 April 2021*
Accepted: *13 July 2021*
Published: *18 August 2021*

Citation:
Sugahara F, Pascual-Anaya J, Kuraku S, Kuratani S and Murakami Y (2021) Genetic Mechanism for the Cyclostome Cerebellar Neurons Reveals Early Evolution of the Vertebrate Cerebellum. Front. Cell Dev. Biol. 9:700860. doi: 10.3389/fcell.2021.700860

The vertebrate cerebellum arises at the dorsal part of rhombomere 1, induced by signals from the isthmic organizer. Two major cerebellar neuronal subtypes, granule cells (excitatory) and Purkinje cells (inhibitory), are generated from the anterior rhombic lip and the ventricular zone, respectively. This regionalization and the way it develops are shared in all extant jawed vertebrates (gnathostomes). However, very little is known about early evolution of the cerebellum. The lamprey, an extant jawless vertebrate lineage or cyclostome, possesses an undifferentiated, plate-like cerebellum, whereas the hagfish, another cyclostome lineage, is thought to lack a cerebellum proper. In this study, we found that hagfish *Atoh1* and *Wnt1* genes are co-expressed in the rhombic lip, and *Ptf1a* is expressed ventrally to them, confirming the existence of r1's rhombic lip and the ventricular zone in cyclostomes. In later stages, lamprey *Atoh1* is downregulated in the posterior r1, in which the *NeuroD* increases, similar to the differentiation process of cerebellar granule cells in gnathostomes. Also, a continuous *Atoh1*-positive domain in the rostral r1 is reminiscent of the primordium of valvula cerebelli of ray-finned fishes. Lastly, we detected a *GAD*-positive domain adjacent to the *Ptf1a*-positive ventricular zone in lampreys, suggesting that the *Ptf1a*-positive cells differentiate into some GABAergic inhibitory neurons such as Purkinje and other inhibitory neurons like in gnathostomes. Altogether, we conclude that the ancestral genetic programs for the formation of a distinct cerebellum were established in the last common ancestor of vertebrates.

Keywords: cerebellum evolution, cyclostome, lamprey, hagfish, evolutionary developmental biology (EvoDevo), rhombic lip, Purkinje cells, granule cells of cerebellum

INTRODUCTION

During development, the cerebellum arises from the dorsal part of rhombomere 1 (r1), just posterior to the midbrain-hindbrain boundary (MHB or isthmic organizer). R1 is defined as an *Otx*- and *Hox*-negative region and this molecular state is critical for cerebellar development (Broccoli et al., 1999; Katahira et al., 2000; Butts et al., 2014a; **Figure 1B**). In addition, FGF8, a secreted molecular signal from the MHB, plays an important role for the induction of the cerebellum (Reifers et al., 1998; Sato et al., 2001). *Wnt1,* which is expressed in the most caudal part of the midbrain, is also known to encode an important signaling molecule. Although not directly involved in the induction of the cerebellum, *Wnt1*-knockout mice lack the cerebellum. This, however, might be a secondary effect as a result of the disruption of the specification of the midbrain and hindbrain (McMahon and Bradley, 1990).

Excitatory granule neurons (glutamatergic), and inhibitory Purkinje cells (GABAergic) play a major role in the processing of the cerebellar cortex. These cerebellar neurons arise from two distinct embryonic domains: the rhombic lip and the ventricular zone (**Figures 1A–A"**). The rhombic lip is the rostrodorsal edge of the hindbrain formed just below the roof plate (Wullimann et al., 2011). The rhombic lip can be divided into the anterior part that arises from r1 and the rest, the posterior part comprised between r2 and 8. In gnathostomes, *Wnt1* and *Pax6* are co-expressed throughout the rhombic lip. Also, *Atoh1* (a bHLH transcription factor; *Math1* in mice)-positive neuron precursors specifically arise from the upper (anterior) rhombic lip, differentiating into cerebellar granule cells, while the cells in the lower (posterior) rhombic lip differentiate into excitatory neurons in the cochlear nuclei and the precerebellar nuclei and give rise to mossy fibers (Fujiyama et al., 2009). In amniotes, *Atoh1*-positive precursors migrate secondarily from the upper rhombic lip to cover the surface of the cerebellar cortex, forming the external granule cell layer (EGL) (Wullimann et al., 2011; **Figures 1A–A"**). The proliferative activity of EGL cells is stimulated transiently by Shh signaling from Purkinje cells (Martí and Bovolenta, 2002). Subsequently, the EGL cells start to express a proneural gene, *NeuroD1,* and stop expressing *Atoh1*, then shifting toward a postmitotic state. Finally, the granule cells migrate ventrally to generate an internal granular layer (Wullimann et al., 2011). This proliferative EGL may have been acquired in the amniote lineage, because frogs have only non-proliferative EGL (Gona, 1972; Butts et al., 2014b) and no EGL has been identified in zebrafish, paddlefish (*Polyodon spathula*), and sharks (*Scylliorhinus canicula*) (Chaplin et al., 2010; Butts et al., 2014c; Hibi et al., 2017). However, there are reports of cell migration and genetic mechanisms similar to the amniote EGL in zebrafish (Biechl et al., 2016), although this remains controversial. The internal migration of the granule cells is also reported in frogs and teleosts, but not in sharks and paddlefish (Chaplin et al., 2010; Butts et al., 2014c; Hibi et al., 2017).

The ventricular zone, another source of cerebellar neurons, occupies the ventromedial position to the rhombic lip (**Figure 1A"**). This domain is characterized by the expression of the bHLH transcription factor coding gene *Ptf1a*, and gives rise to GABAergic inhibitory neurons including Purkinje, basket, stellate, and Golgi cells in the cerebellar cortex (Leto et al., 2016). A *Ptf1a* knockout mouse, *cerebelles*, lacks the entire cerebellar cortex together with Purkinje cells and other GABAergic interneurons, indicating that *Ptf1a* is essential for the development of GABAergic inhibitory neurons in the cerebellum (Hoshino et al., 2005; Hoshino, 2006). The caudal area (r2-r8) of the ventricular zone produces inhibitory neurons in the cochlear nuclei and neurons that generate climbing fibers in the precerebellar nuclei (Yamada et al., 2007; Fujiyama et al., 2009).

Cyclostomes that include lampreys and hagfish, are the only living jawless vertebrates and diverged from gnathostomes over 500 million years ago (Kuraku and Kuratani, 2006). The lamprey possesses only a small, plate-like cerebellum in the rostrodorsal part of the hindbrain, in which small granule-like cells are present (Striedter and Northcutt, 2020). Therefore, the lamprey cerebellum is considered to be homologous to the corpus cerebelli (the central lobe of the cerebellum) of gnathostomes (Nieuwenhuys, 1967). However, the lamprey lacks a layered cerebellar cortex. The presence of the Purkinje cells is still ambiguous in lamprey (Lannoo and Hawkes, 1997). On the other hand, the hagfish, the other group of cyclostomes, does not possess a morphologically distinct cerebellum, showing only the vestibulo-lateral commissure reported at the midbrain-hindbrain boundary (Larsell, 1947; Nieuwenhuys et al., 1998; Sugahara et al., 2016).

Previous studies have shown that lamprey embryos exhibit *Otx*- and *Hox1*-negative rhombomere 1 region (Ueki et al., 1998; Takio et al., 2004), while *Gbx* is expressed in the entire hindbrain (Takio et al., 2007). As in gnathostomes, the lamprey MHB expresses *FGF8* and *Wnt1* (Shigetani et al., 2002; Sugahara et al., 2016; **Supplementary Figure 1**), suggesting that a certain regionalization and inductive signals for the cerebellum are shared at least by lampreys. Regarding the development of cerebellar neurons, we previously reported that *Atoh1* is expressed in the lamprey in dorsal rhombomeres, including r1, and one of two orthologous genes of *Ptf1a* expressed in the ventral part (Sugahara et al., 2016). However, it is still unclear whether cerebellar neuron subtypes are differentiated from these domains or not. Much less is known about the presence or absence of this regionalization in the hagfish, the involved induction signals or the genetic program for the specification of cerebellar neurons subtypes. We previously reported that *Pax6* and *Atoh1* are expressed in the rhombic lip of the hagfish but could not identify exactly if it was in r1 only or also in more caudal rhombomeres, because we could use only sectioned samples due to the limitation of embryonic material (Ota et al., 2007). Here, we further investigated the molecular basis underlying the cerebellum regionalization and the cerebellar neuron specification in lamprey and hagfish embryos. Also, by comparing gnathostomes and cyclostomes, we tried to depict the molecular and morphological ground patterns of the cerebellum of the last common ancestor of crown vertebrates, and to infer the evolutionary changes in its developmental mechanisms in the different vertebrate lineages.

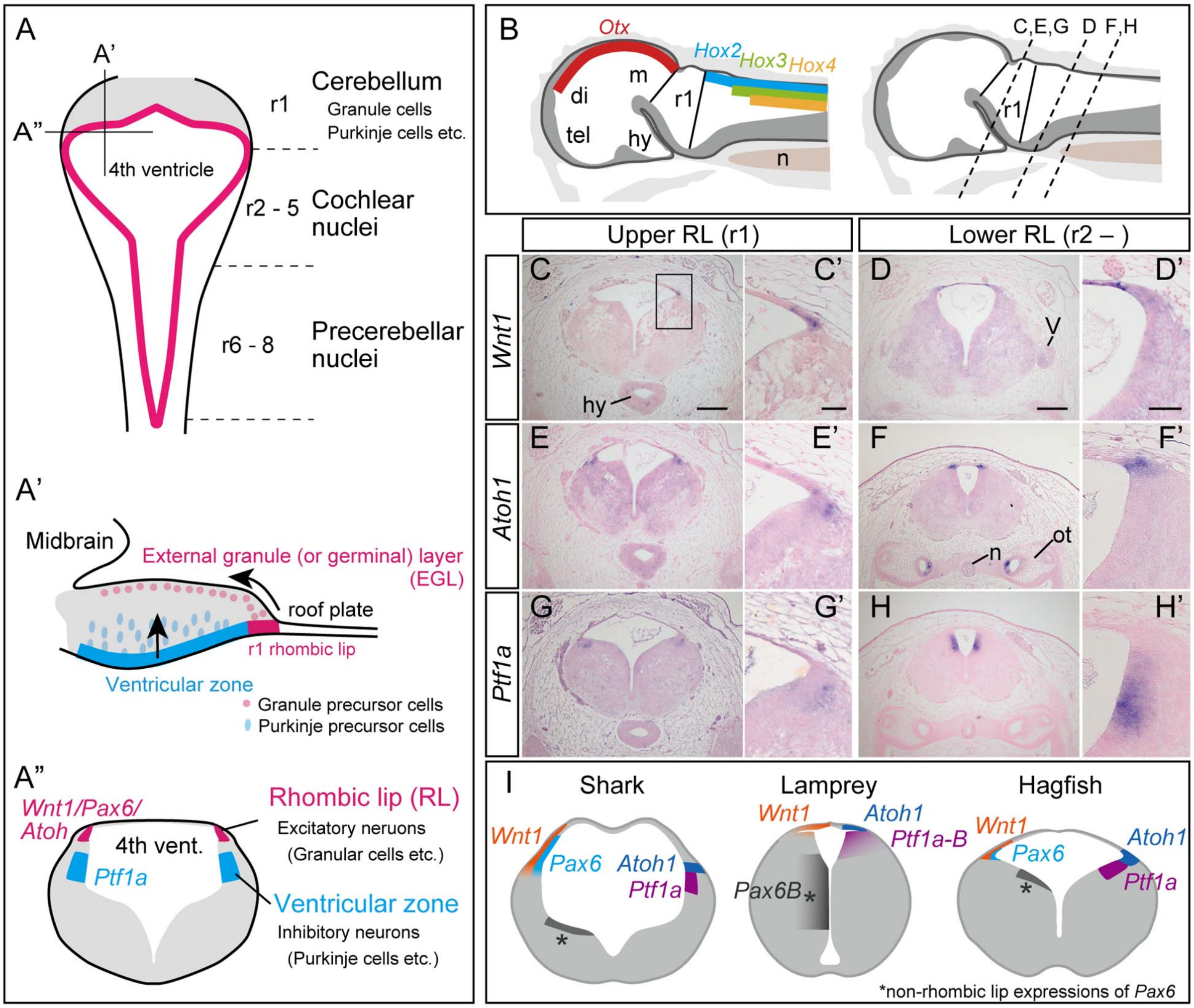

A
A'
A"
Cerebellum
Granule cells
Purkinje cells etc.
r1
4th ventricle
r2 - 5
Cochlear nuclei
r6 - 8
Precerebellar nuclei
A'
Midbrain
External granule (or germinal) layer (EGL)
roof plate
r1 rhombic lip
Ventricular zone
Granule precursor cells
Purkinje precursor cells
A"
Wnt1/Pax6/Atoh
Rhombic lip (RL)
Excitatory neruons
(Granular cells etc.)
4th vent.
Ptf1a
Ventricular zone
Inhibitory neurons
(Purkinje cells etc.)
B
Otx
Hox2
Hox3
Hox4
m
di
tel
hy
r1
n
C,E,G D F,H
r1
Upper RL (r1)
Lower RL (r2 –)
Wnt1
C
C'
D
D'
V
hy
Atoh1
E
E'
F
F'
n
ot
Ptf1a
G
G'
H
H'
I
Shark
Lamprey
Hagfish
Wnt1
Pax6
Atoh1
Ptf1a
Wnt1
Atoh1
Ptf1a-B
Pax6B
Wnt1
Pax6
Atoh1
Ptf1a
*non-rhombic lip expressions of Pax6

FIGURE 1 | Development of the cerebellum in mammals, based on Fujiyama et al., 2009 **(A)**. **(A)** Dorsal view of the rhombomeres. The rhombomere 1 (r1) gives rise to the cerebellum, while r2–8 generates the cochlear nuclei and precerebellar nuclei. **(A')** Sagittal view of the rhombomere 1. Granule precursor cells, generated in the r1 rhombic lip, migrate tangentially to form the external germinal layer (EGL). The Purkinje precursor cells originating from the ventricular zone secrete Shh, which is involved in the proliferation of the granule cells. The granule cells in EGL then migrate ventrally to form the granular layer. **(A")** Transverse view of rhombomere 1 showing the rhombic lip as the source for excitatory neurons, and the ventricular zone where the inhibitory neurons arise. **(B–H)** Upper and lower rhombic lip gene expression patterns in the hagfish, *E. burgeri*. **(B)** Illustration of a sagittal section of the anterior part of a stage 53 embryo. (left) Rhombomere 1 (Upper rhombic lip) is identified as a negative region of both *Otx* and *Hox* genes (Oisi et al., 2013; Pascual-Anaya et al., 2018). (right) Anterior-posterior level of transverse sections shown in **(C–H)** are indicated in a sagittal illustration. **(C–H)** Section *in situ* hybridizations for hagfish *Wnt1*, *Atoh1* and *Ptf1a* genes showing their expression patterns in r1 **(C,E,G)** and posterior to r1 **(D,F,H)**. **(C'–H')** Higher magnification images of the rhombic lip expression patterns shown in **(C–H)**. **(I)** Comparison of the upper rhombic lip gene expression patterns in the shark (*S. torazame*), lamprey (*L. camtschaticum*), and hagfish (*E. burgeri*) based on Sugahara et al. (2017) and the present study. di, diencephalon; hy, hypothalamus; m, mesencephalon; n, notochord; ot, otic capsule; tel, telencephalon; r1, rhombomere 1; V, trigeminal nerve. Scale bars: 200 μm **(C–H)**, 50 μm **(C'–H')**.

MATERIALS AND METHODS

Animals

Matured arctic lamprey, *Lethenteron camtschaticum* were collected in Hokkaido, Japan. Embryos were obtained by artificial fertilization and staged as described previously (Tahara, 1988; Sugahara et al., 2015) and fixed with 4% paraformaldehyde. Matured Japanese inshore hagfish, *Eptatretus burgeri* were caught from the Japan Sea off Shimane Prefecture. The embryos of *E. burgeri* were obtained by keeping adults in cages and placed in the sea bottom as described previously (Oisi et al., 2015). The staging of hagfish embryos followed Dean (1899) and Oisi et al. (2013) and fixed with Serra's fixative (6:3:1 mixture of ethanol, formalin, and glacial acetic acid).

Gene Identification and Phylogenetic Analysis

Sequence of hagfish *Wnt1* and *Ptf1a* genes were obtained from a previously published transcriptome assembly (Pascual-Anaya et al., 2018) by means of TBLASTN v2.2.31 + searches

(Altschul et al., 1997) and using gnathostome counterparts as queries. *Ptf1a* sequence was completed on the 3′ end by manual cloning and sequence following standard protocols. For the lamprey *GAD* gene, the sequence was obtained from transcriptome data reported previously (Pascual-Anaya et al., 2018), based on the sequence of *Petromyzon marinus GAD1* (XM_032965485). Reverse transcription polymerase chain reaction (RT–PCR) was performed to amplify fragments of each gene with specific primers from total RNA of lamprey embryos or hagfish juvenile (sequences of primers used are listed in **Supplementary Table 1**). The sequences identified here have been deposited in GenBank under accession numbers LC60429-31. Hagfish *Atoh1* (KT897935), lamprey *Atoh1* (KT897930), *Ptf1a-B* (KT897934), *NeuroD* (LC424505) were reported previously (Sugahara et al., 2016; Higuchi et al., 2019).

Molecular phylogenetic tree inference of the *Wnt1* and *Ptf1a* genes were performed as previously described (Hara et al., 2018) using homolog sequences retrieved via the aLeaves webserver (Kuraku et al., 2013).

In situ Hybridization of the Hagfish and Lamprey Embryos

In situ hybridization on paraffin sections of *E. burgeri* was performed as described (Oisi et al., 2015). Whole-mount *in situ* hybridization on lamprey embryos was performed according to Sugahara et al., 2015. Some stained specimens were made transparent by 1:2 mixture of Benzyl-alcohol and Benzyl-benzoate after dehydration. Staining specimens were embedded in tissue tek® O.C.T.^{TM} compound and sectioned (12 μm) by CryoStar NX50 (Thermo Fisher Scientific).

RESULTS

Isthmic Gene Expressions in Cyclostomes

Previously, we have identified the expression of *FGF8/17* and *Wnt1* in a sharp domain, of the lamprey brain and corresponding to the MHB (Shigetani et al., 2002; Sugahara et al., 2016; **Supplementary Figures 1B,D**). However, the presence of this expression domain in the hagfish is so far unknown. We newly identified a hagfish *Wnt1* orthologous gene (**Supplementary Figure 3A**) and investigated its expression pattern in the hagfish brain. At mid-pharyngula stage (Dean stage 45), hagfish *Wnt1* expression was observed as a sharp transverse band corresponding to the MHB (**Supplementary Figure 1A**). Also, hagfish *FGF8/17* was expressed around the MHB, but in a relatively broader domain than that of lamprey and gnathostomes (**Supplementary Figure 1C**).

Expression Patterns of Upper- and Lower-Rhombic Lip Marker Genes in Hagfish

Next, we investigated whether the genetic machinery for r1-neuronal differentiation is present in cyclostomes or not. Generally, when looked from a dorsal view, the gnathostome hindbrain appears as a rhomboid shape with a thin roof plate over the fourth ventricle. The most expanded region corresponds to the boundary between r1 and r2 (Wingate, 2001). In hagfish, the anterior border of the *Hox2* expression was seen in the lateral edge of the fourth ventricle (Pascual-Anaya et al., 2018). We therefore defined this region as the posterior limit of the anterior rhombic lip (**Figure 1B**). At late pharyngular stage (Dean stage 53), *Atoh1* was expressed in the bilateral dorsal-most edges of the hindbrain, the "rhombic lip," in the upper and lower rhombomeres. *Wnt1* expression overlaps with *Atoh1* in the rhombic lip but extends to the roof plate (**Figures 1C–F'**). We newly identified a *Ptf1a* ortholog gene in hagfish (**Supplementary Figure 3B**). Hagfish *Ptf1a* is expressed at the surface of the fourth ventricle ventral to the *Atoh1*/*Wnt1* expressing region in both lower and upper rhombomeres (**Figures 1G,H**). These results indicate that the regionalization of the rhombic lip and the ventricular zone are present in hagfish r1 (**Figure 1I**).

Spatio-Temporal Atoh1 Expression in the Lamprey Hindbrain

In zebrafish and paddlefish, granule neuron precursors are reported to express *Atoh1* in early embryonic stages, but *Atoh1* is later downregulated in the lateral cerebellar rhombic lip. Subsequently, post-mitotic granule precursor cells express *NeuroD* and differentiate into matured granule cells. Meanwhile, proliferation of the precursor cells, which express *Atoh1*, is selectively maintained in the rostromedial rhombic lip until late development (Kani et al., 2010; Butts et al., 2014c). To investigate whether the same or a similar genetic mechanism for the differentiation of the cerebellar granule cells is present in the lamprey or not, we observed the spatio-temporal expression pattern of *Atoh1* in lamprey embryos. *Atoh1* expression starts at stage 23.5 in the rostral half of the rhombic lip (**Figure 2B**) and then expands throughout the upper and lower rhombic lip during stages 24–26 (**Figures 2C–E**). Subsequently, this *Atoh1* expression is downregulated in the caudal half of r1 and lower rhombic lip, whereas its expression is maintained in the rostral tip of r1. This continuous expression of *Atoh1* is observed at least until respiration stage (stage 28) (**Figure 2F**). Like ray-finned fish, the rostral area of lamprey r1 is apparently regulated differently from the other rhombic lip (Kani et al., 2010; Butts et al., 2014c). In birds and mammals, *Atoh1*-positive granule precursor cells migrate from the rhombic lip to cover the surface of the cerebellum to form the EGL. Subsequently, EGL cells migrate ventrally and form the internal granule layer (Butts et al., 2014a). In the lamprey, however, we could not observe the expansion or migration of *Atoh1*-positive cells as is the case in the shark and paddlefish (Chaplin et al., 2010; Butts et al., 2014c).

Differentiation of Atoh1–Positive Cells in the Lamprey Late Development

As mentioned above, lamprey *Atoh1* is downregulated at late stages except in the rostral area of the r1 (**Figure 2F**). In

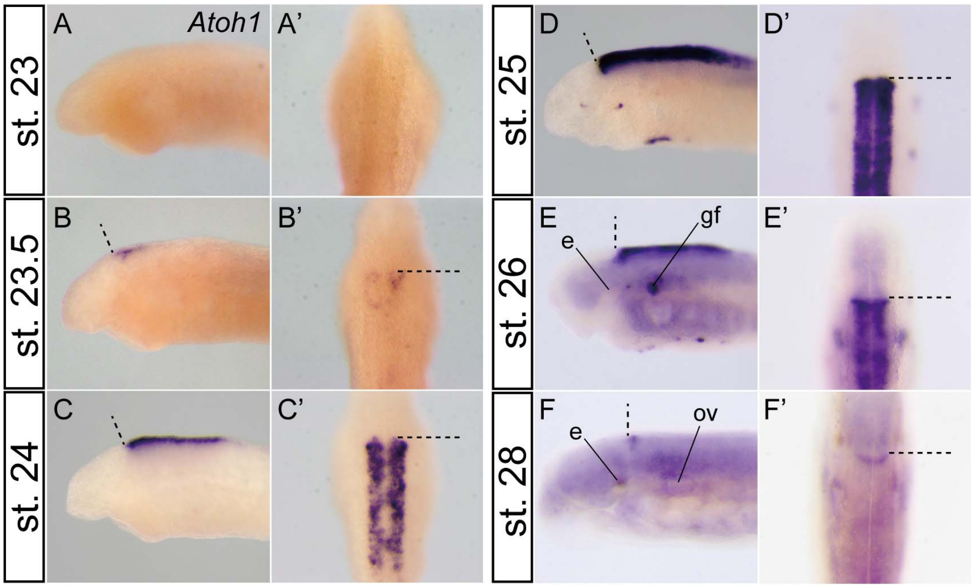


FIGURE 2 | Temporal *Atoh1* expression pattern in the lamprey hindbrain. Lateral **(A–F)** and dorsal views **(A'–F')**. Dashed lines indicate the MHB. *Atoh1* expression starts at stage 23.5 in the rostral half of the rhombic lip **(B)** and subsequently expands throughout the rhombic lip during stages 24–26 **(C–E)**. After stage 27 (not shown), the expression is restricted only to the rostral border of the rhombomere 1 **(F)**. Note that the apparent asymmetrical staining in **(B)** does not reflect the actual difference in expression, but is caused by the staining process. e, eye; gf, complex of facial/lateral line ganglion; ov, otic vesicle.

accordance with the decrease of *Atoh1* expression, the *NeuroD* expression become apparent slightly caudal to the *Atoh1*-positive area (**Figures 3A,B** and **Supplementary Figure 2**). To distinguish whether the *NeuroD*-positive region is included within r1 or not, we conducted double *in situ* hybridization with *OtxA* and *Hox2α* (**Figures 3C,D**). Since *OtxA* and *Hox2α* are expressed in the midbrain and r2, respectively (Ueki et al., 1998; Takio et al., 2004), we could identify r1 as the *OtxA*/*Hox2α*-negative region in between. By comparing *NeuroD* expression with *Otx*–*Hox2* expression, we identified the *NeuroD*–positive region located in the posterior part of r1 (**Figure 3E**). Transverse sections showed that *NeuroD* expression was located lateral to, but slightly overlapping the *Atoh1* expression domain (**Figures 3A',B'**). The topological location of these two *Atoh1* and *NeuroD* expression domains suggests that precursors migrate from more anterior/medial (*Atoh1*-positive) areas to more posterior/lateral (*NeuroD*-positive) zones during their differentiation process. At the same time, however, neuron precursors in the rostral r1 retain *Atoh1* expression. This positional relationship is comparable to the development of cerebellar granule neurons in some gnathostome fishes as discussed below.

Ptf1a Expression and GAD–Positive Inhibitory Neurons

In mammals, *Ptf1a*–positive cells in r1 ventricular zone differentiate into GABA-releasing inhibitory neurons such as Purkinje cells and several other types of interneurons (basket, stellate, and Golgi cells) (Hoshino et al., 2005). Presence of *Ptf1a*–positive cells in r1 ventricular zone in both lampreys and hagfish suggests that an active and probably homologous genetic mechanism for GABAergic inhibitory neuron production exist in cyclostomes. We therefore observed the expression of *glutamate decarboxylase* (*GAD*) gene, which encodes an enzyme to catalyze the synthesis of GABA. In lamprey, at stage 26, *GAD* expression domain overlaps the *Ptf1a* expression domain in the anterior part of the rhombencephalon (including r1) (arrowheads in **Figures 3F,G**). Transverse sections showed that *GAD* expression localizes relatively lateral to the *Ptf1a*–positive ventricular zone (**Figures 3F',G'**). Since *Ptf1a* is expressed only in undifferentiated neurons in the neuroepithelium (Hoshino et al., 2005), this may possibly reflect that *Ptf1a*-positive precursor cells in lamprey are generated in the ventricular zone, then migrate radially, and finally differentiate into inhibitory GABAergic neurons.

DISCUSSION

Presence of MHB Signals in Cyclostomes

Signaling factors from the MHB are essential for the induction and development of the gnathostome cerebellum (Leto et al., 2016). In this study, we identified both *FGF8/17* and *Wnt1* expression around the MHB region in the hagfish embryo similar to gnathostomes and the lamprey (**Supplementary Figure 1**), suggesting that these signals are conserved through all extant vertebrates. Curiously, hagfish *FGF8/17* expression showed a broader pattern than in other vertebrates (**Supplementary Figure 1C**). This unique

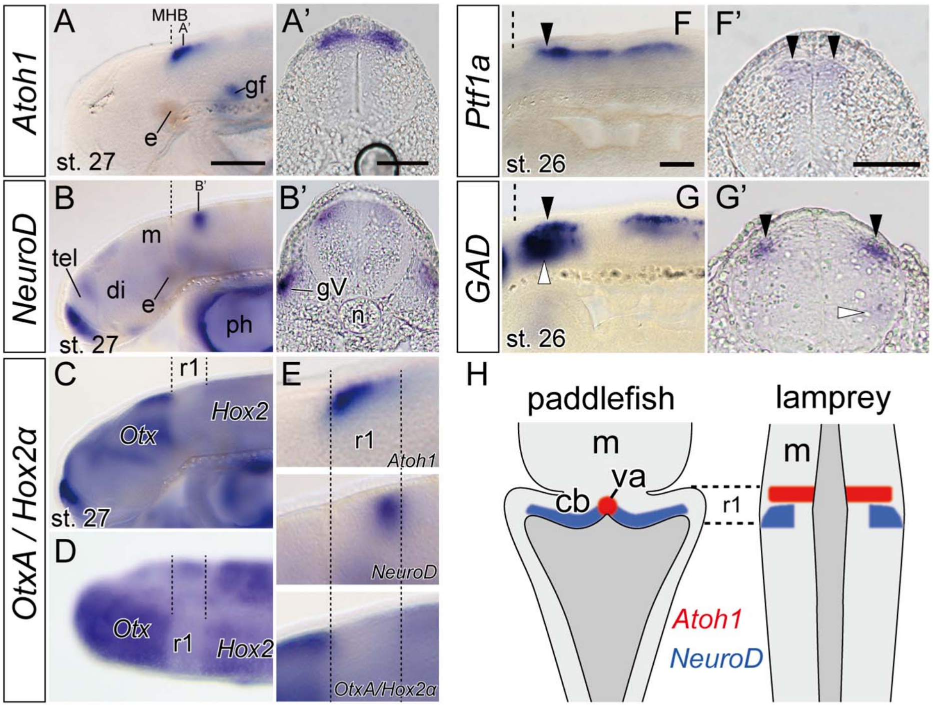


FIGURE 3 | Expression of genes involved in neuron differentiation during lamprey late embryonic stages. **(A,A')** *Atoh1* expression is restricted to the anterior r1 in stage 27 embryo (arrow heads). Dashed lines indicate the MHB. **(B,B')** *NeuroD* expression is observed behind the *Atoh1* expression domain (arrowheads). Note that **(A',B')** are not at the same rostrocaudal positions in r1. The apparent asymmetrical staining in **(B')** does not reflect the actual difference in expression, but is caused by the slightly oblique cutting of the sections. **(C,D)** *In situ* hybridization of both *OtxA* and *Hox2α* probes in stage 27 in lateral view **(C)** and dorsal view **(D)**. *OtxA* is expressed only anterior to the midbrain, whereas *Hox2α* is expressed caudally from r2. Dashed lines indicate the borders of r1 identified as "Otx–Hox free" domain. **(E)** High magnifications of the panel **(A–C)** in r1 region. **(F,F')** *Ptf1a-B* expression at the ventricular zone at stage 26 (arrowheads). **(G,G')** *GAD* is expressed in the hindbrain at the same dorsal level than –but lateral to– the *Ptf1a* expression domains (black arrowheads) and also in a ventral domain (white arrowheads). **(H)** Comparison of *Atoh1* and *NeuroD* expression patterns in the hindbrain between paddlefish, *Polyodon spathula* (stage 42, based on Butts et al., 2014c), and lamprey. Note that although the lamprey hindbrain does not show a 'diamond shape' of the rhombic lip during embryonic stages, the topological relationship of these gene expression patterns is comparable in both species. cb, cerebellum primordium; gV, trigeminal ganglion; ph, pharynx; va, valvula cerebelli. See **Figures 1**,**2** for other abbreviations. Scale bars, 100 μm **(A–D)**, 50 μm **(A',B',F,F',G,G')**.

FGF8/17 expression pattern in the hagfish might be related to the absence of a distinct cerebellum in adult hagfish. Alternatively, FGF8 signaling might have a limited effect on cerebellar development, since some reports have indicated that conditional mutant mice with reduced *FGF8* expression in the MHB shows only loss of the cerebellar vermis region (Basson et al., 2008; Yu et al., 2011, 2013). Further experiments would be needed to explain this unusual *FGF8/17* expression pattern in the hagfish.

Upper and Lower Rhombic Lip and Ventricular Zone in Cyclostomes

We morphologically defined the domain of r1 in paraffin sections of a hagfish embryo and found that *Wnt1*and *Atoh1* are expressed not only in r1, but also in the more caudal rhombomeres. Expression of *Pax6*, another rhombic lip gene, was previously reported (see Figure 4C in Sugahara et al., 2016). In addition, we identified *Ptf1a* expression in the ventral side of the *Wnt1/Atoh1* expression domain (**Figures 1G,H**). Together with previous studies in the lamprey, we conclude that the regionalization into the upper/lower rhombic lip and the accompanying ventricular zone is shared between cyclostomes and gnathostomes (**Figure 1I**).

Presence of a Differentiation Mechanism for Granule Cells and a Continuous *Atoh1*-Expression Domain

Lamprey *Atoh1* is initially expressed in the entire rhombic lip, but subsequently gets downregulated in the caudal half of r1 and the lower rhombic lip (**Figure 2**). Concomitantly with this downregulation, *NeuroD* becomes upregulated in the caudal part of r1 (**Supplementary Figure 2** and **Figure 3B**). This transition is consistent with the differentiation process of granule cells in zebrafish, paddlefish and amniotes (Kani et al., 2010; Butts et al., 2014c). Therefore, the genetic mechanism for the differentiation of the cerebellar granule cells might be present in the lamprey as well.

We found a continuous expression domain of the *Atoh1* in the rostral r1, topologically comparable to the rostromedial domain of the paddlefish and zebrafish r1 (Kani et al., 2010;

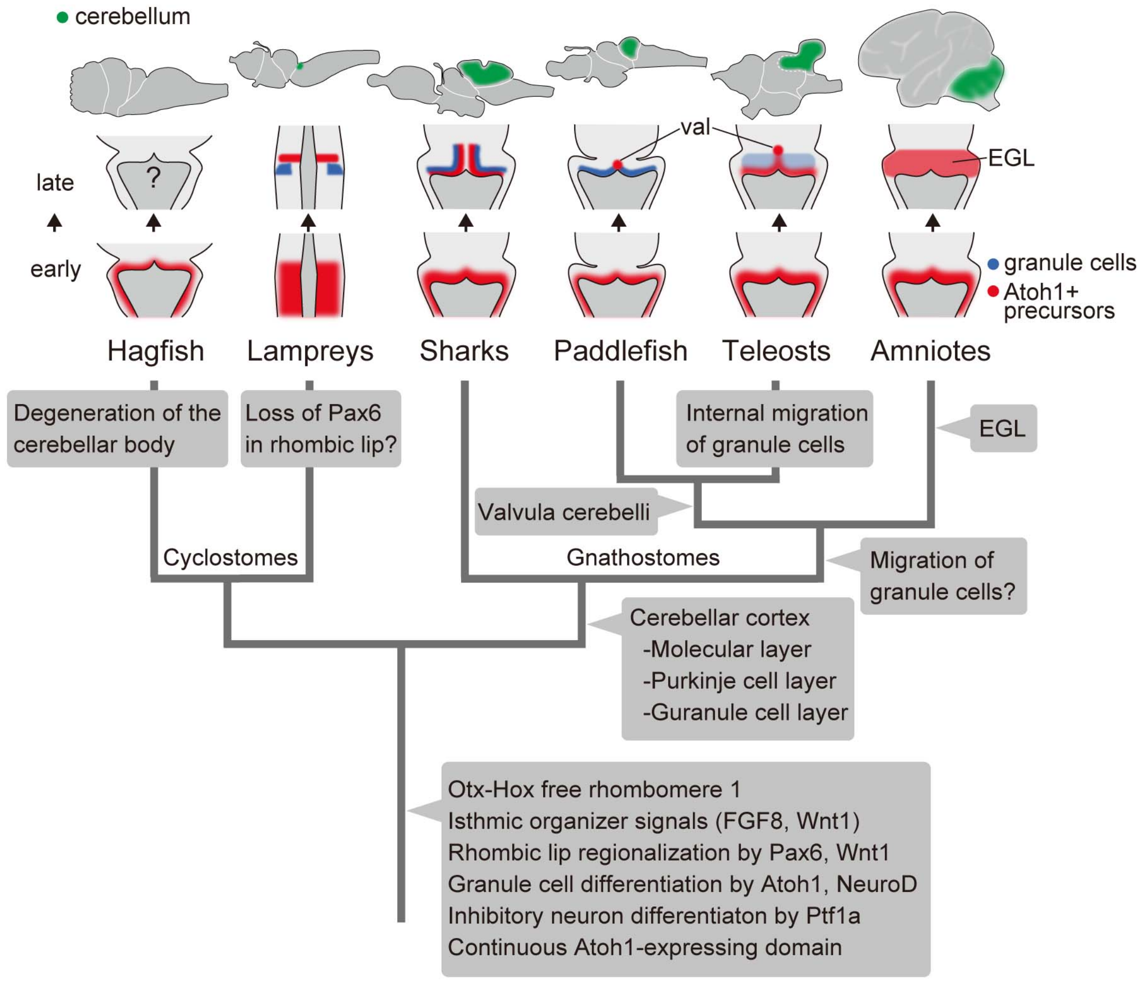

● cerebellum
val
EGL
late
?
early
● granule cells
● Atoh1+ precursors
Hagfish
Lampreys
Sharks
Paddlefish
Teleosts
Amniotes
Degeneration of the cerebellar body
Loss of Pax6 in rhombic lip?
Internal migration of granule cells
EGL
Valvula cerebelli
Cyclostomes
Gnathostomes
Migration of granule cells?
Cerebellar cortex
-Molecular layer
-Purkinje cell layer
-Guranule cell layer
Otx-Hox free rhombomere 1
Isthmic organizer signals (FGF8, Wnt1)
Rhombic lip regionalization by Pax6, Wnt1
Granule cell differentiation by Atoh1, NeuroD
Inhibitory neuron differentiaton by Ptf1a
Continuous Atoh1-expressing domain

FIGURE 4 | An evolutionary scenario for the vertebrate cerebellum, based on Chaplin et al. (2010) and Butts et al. (2014c) and the present study. Note that the internal migration of postmitotic granule cells has been observed in teleosts but not in the paddlefish. Therefore, it is equally possible that this migration of granule cells was acquired after the divergence of chondrichthyans and osteichthyans, with paddlefish losing this ability secondarily. Alternatively, it is possible that teleosts acquired the migration independently from amniotes. val, valvula cerebelli; EGL, external granule cell layer.

Butts et al., 2014c; **Figure 3H**). In the zebrafish, this region differentiates into valvula cerebelli, a unique cerebellar structure in ray-finned fishes but not in cartilaginous fish and lobe-finned fish including tetrapods (Nieuwenhuys et al., 1998). Although the valvula cerebelli has not been identified anatomically in adult lampreys, it is possible that a homologous genetic domain might be present in the embryonic lamprey r1.

Presence of GAD-Positive Inhibitory Neurons in Lamprey r1

In mice, all cerebellar inhibitory neurons arise from the *Ptfa1*-positive ventricular zone (Hoshino et al., 2005). Also, inhibitory neurons in the cochlear nucleus arise from r2–5, and the climbing fibers originate from the inferior olivary nucleus arise from r6–8 (Yamada et al., 2007; Fujiyama et al., 2009). In this study, orthologous *Ptf1a* genes were expressed in the ventricular zone in lamprey and hagfish r1 and more posterior rhombomeres (**Figures 1**, **3**). In addition, lamprey *GAD and Ptf1a* expression domains were observed at the same anterior-posterior level in the rhombomeres with a distinct dorso-ventral pattern (**Figure 3G**). If *GAD*-positive inhibitory neurons are derived from the *Ptf1a*-positive ventricular zone, GABAergic inhibitory neurons in the lamprey may be formed by a similar developmental mechanism to that of gnathostome Purkinje cells. However, so far, Purkinje cells have not been identified in the adult lamprey cerebellum (Nieuwenhuys et al., 1998; Striedter and Northcutt, 2020). Further studies are necessary to determine whether these GABAergic neuron precursors in the lamprey embryo are homologous to Purkinje cells or any other inhibitory interneurons in gnathostomes or not (e.g., basket, stellate, and Golgi cells). In mice, small GABAergic neurons in the deep cerebellar nuclei (DCN) arise from the *Ptf1a*-positive ventricular zone and migrate ventrally (Hoshino et al., 2005). It is necessary to confirm whether the ventral population of *GAD*-expressing cells in the lamprey (white arrowheads in **Figure 3G**) originate from the ventricular zone and migrate ventrally like DCN neurons in amniotes. In mice, a subpopulation of *Ptf1a*-positive cells in the ventricular zone expressing *Gsx1* differentiates into interneurons in the DCN (Seto et al., 2014). Two *Gsx1* homolog genes have been identified in the lamprey and one of them was amplified from the brain tissue by RT-PCR (Zhang et al., 2017). However, further research is needed to determine whether a similar neurodevelopmental mechanism for the differentiation of *Ptf1a*-expressing cells in the ventricular zone is present in the lamprey.

Developmental Evolution in the Vertebrate Cerebellum

By comparing both cyclostomes and gnathostomes, we can depict a hypothetical evolutionary scenario of the vertebrate cerebellum, as shown in **Figure 4**. The developing brain of the last common ancestor of extant vertebrates already had *Otx-Hox* free r1, with isthmic organizer signals such as FGF8 and Wnt1. Given the presence of similar expression field with conserved function in the hemichordate larvae, the isthmic FGF signaling may date back to deep deuterostome ancestry (Pani et al., 2012). In the r1 rhombic lip, *Pax6* and *Wnt1* specify a domain, from which the granule cells differentiate as the result of *Atoh1* and *NeuroD* expression. Also, *Ptf1a*-positive ventricular zone generates some GABAergic inhibitory neurons.

We can also trace lineage-specific evolutionary modifications to this ancestral state of the cerebellum. In the hagfish, although the rhombic lip and ventricular zone are present in the embryo, most structures and neurons were degenerated developmentally in the adult brain.

Unlike in gnathostomes and hagfish, lamprey *Pax6a* and *6b* are not expressed in the rhombic lip (Murakami et al., 2001; Sugahara et al., 2016). In *Pax6* knockout mice, proliferation and initial differentiation of granule cells are unaffected, while cell migration to form the EGL is disrupted (Engelkamp et al., 1999). So far, migration of granule cells has not been observed in the lamprey as in cartilaginous fish and paddle fish. It is possible that the lamprey lost this *Pax6* expression domain in the rhombic lip secondarily. Subsequently, *Pax6* may have become involved in the migration of granule cells and the formation of the EGL at a certain period of gnathostome evolution (**Figure 4**). Recently, a third lamprey *Pax6* gene was identified (Ravi et al., 2019). Future experiments will be necessary to determine whether this *Pax6* paralog is expressed in the lamprey rhombic lip or not, which eventually could change this evolutionary scenario (**Figure 4**).

The cerebellar cortex was established in gnathostomes before the divergence of chondrichthyans and osteichthyans. At this period, Purkinje cells started expressing *Shh*, which is not seen in the lampreys (Sugahara et al., 2011). Finally, the proliferative activity of EGL cells by SHH signaling was acquired in the course of amniote evolution. In conclusion, we suggest that the *Atoh1*/*NeuroD* axis for granule cell fate and *Ptf1a* role in inhibitory neuronal specification are genetic programs involved in cerebellar development that were already established in the last common ancestor of living vertebrates. Subsequently, the developmental program of the cerebellum has evolved to adapt to complex environments especially in gnathostomes, presumably related to the acquisition of balancing organs such as paired appendages and inner ear with three semicircular canals.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Experiment Committee of Hyogo College of Medicine.

AUTHOR CONTRIBUTIONS

FS, JP-A, SKurat, and YM conceived the project, designed the experiments, and wrote the manuscript. FS performed animal sampling and gene cloning and gene expression analyses. JP-A searched gene sequences from NGS databases. SKurak performed the gene phylogenetic analysis. All authors analyzed and discussed the data, and approved the final version of the manuscript.

FUNDING

This project has been supported by the RIKEN, Kobe, Japan, the Japan Society for the Promotion of Science (JSPS; grant no. 19K06794 to FS, 20K06892 to YM) and Brain Science Foundation.

ACKNOWLEDGMENTS

We thank O. Kakitani, K. Ota, Y. Oisi, W. Takagi, and S. Higuchi for hagfish sampling; R. Kusakabe, Noboru Sato, H. Nagashima, and Naohiro Sato for lamprey sampling; S. Shibuya, K. Yamamoto, and Y. Yamamoto for maintenance of hagfish aquarium tanks; M. Kawaguchi for discussions. We are grateful to the past and current members of the Evolutionary Morphology Laboratory at RIKEN for their technical advice, animal sampling and maintenance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.700860/full#supplementary-material

Supplementary Figure 1 | Expression of genes involved signaling pathways of the isthmic organizer. *Wnt1* **(A,B)** and *FGF8/17* **(C,D)** in hagfish stage 45 **(A,C)**

and lamprey stage 26 **(B,D)**. Arrowheads mark the isthmic organizer region in the most caudal part of the midbrain. lv, lateral ventricle; 4v, fourth ventricle; ph, pharynx.

Supplementary Figure 2 | Temporal expression of the *NeuorD* in the rhombomere 1. Dorsal views of the head region of the lamprey stages 25 **(A)**, 26 **(B)**, and 27 **(C)**. Arrowheads in **(B,C)** indicate the expression in the rhombomere 1. gV_1, ophthalmic ganglion; $gV_{2,3}$, maxillomandibular ganglion.

Supplementary Figure 3 | Phylogenetic trees of *Wnt1* **(A)** and *Ptf1a* **(B)** genes. The trees were inferred with the maximum-likelihood method using 355 and 153 aligned amino acid sites, respectively. The support values at nodes indicate bootstrap values and posterior probabilities based on the maximum-likelihood method and Bayesian inference in order, respectively. See the section "Materials and Methods" for details.

Supplementary Table 1 | Primers used for gene cloning and probe synthesis.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Basson, M. A., Echevarria, D., Ahn, C. P., Sudarov, A., Joyner, A. L., Mason, I. J., et al. (2008). Specific regions within the embryonic midbrain and cerebellum require different levels of FGF signaling during development. *Development* 135, 889–898. doi: 10.1242/dev.011569

Biechl, D., Dorigo, A., Koster, R. W., Grothe, B., and Wullimann, M. F. (2016). Eppur Si Muove: evidence for an external granular layer and possibly transit amplification in the teleostean cerebellum. *Front. Neuroanat.* 10:49. doi: 10.3389/fnana.2016.00049

Broccoli, V., Boncinelli, E., and Wurst, W. (1999). The caudal limit of Otx2 expression positions the isthmic organizer. *Nature* 401, 164–168. doi: 10.1038/43670

Butts, T., Green, M. J., and Wingate, R. J. (2014a). Development of the cerebellum: simple steps to make a 'little brain'. *Development* 141, 4031–4041. doi: 10.1242/dev.106559

Butts, T., Hanzel, M., and Wingate, R. J. (2014b). Transit amplification in the amniote cerebellum evolved via a heterochronic shift in NeuroD1 expression. *Development* 141, 2791–2795. doi: 10.1242/dev.101758

Butts, T., Modrell, M. S., Baker, C. V., and Wingate, R. J. (2014c). The evolution of the vertebrate cerebellum: absence of a proliferative external granule layer in a non-teleost ray-finned fish. *Evol. Dev.* 16, 92–100. doi: 10.1111/ede.12067

Chaplin, N., Tendeng, C., and Wingate, R. J. (2010). Absence of an external germinal layer in zebrafish and shark reveals a distinct, anamniote ground plan of cerebellum development. *J. Neurosci.* 30, 3048–3057. doi: 10.1523/jneurosci.6201-09.2010

Dean, B. (1899). *On the Embryology of Bdellostoma stouti. A General Account of Myxinoid Development from the Egg and Segmentation to Hatching. Festschrift zum 70ten Geburststag Carl von Kupffer*. Jena: Verlag von Gustav Fischer in Jena, 220–276.

Engelkamp, D., Rashbass, P., Seawright, A., and van Heyningen, V. (1999). Role of Pax6 in development of the cerebellar system. *Development* 126, 3585–3596.

Fujiyama, T., Yamada, M., Terao, M., Terashima, T., Hioki, H., Inoue, Y. U., et al. (2009). Inhibitory and excitatory subtypes of cochlear nucleus neurons are defined by distinct bHLH transcription factors, Ptf1a and Atoh1. *Development* 136, 2049–2058. doi: 10.1242/dev.033480

Gona, A. G. (1972). Morphogenesis of the cerebellum of the frog tadpole during spontaneous metamorphosis. *J. Comp. Neurol.* 146, 133–142. doi: 10.1002/cne.901460202

Hara, Y., Yamaguchi, K., Onimaru, K., Kadota, M., Koyanagi, M., Keeley, S. D., et al. (2018). Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat. Ecol. Evol.* 2, 1761–1771. doi: 10.1038/s41559-018-0673-5

Hibi, M., Matsuda, K., Takeuchi, M., Shimizu, T., and Murakami, Y. (2017). Evolutionary mechanisms that generate morphology and neural-circuit diversity of the cerebellum. *Dev. Growth Differ.* 59, 228–243. doi: 10.1111/dgd.12349

Higuchi, S., Sugahara, F., Pascual-Anaya, J., Takagi, W., Oisi, Y., and Kuratani, S. (2019). Inner ear development in cyclostomes and evolution of the vertebrate semicircular canals. *Nature* 565, 347–350. doi: 10.1038/s41586-018-0782-y

Hoshino, M. (2006). Molecular machinery governing GABAergic neuron specification in the cerebellum. *Cerebellum* 5, 193–198. doi: 10.1080/14734220600589202

Hoshino, M., Nakamura, S., Mori, K., Kawauchi, T., Terao, M., Nishimura, Y. V., et al. (2005). Ptf1a, a bHLH transcriptional gene, defines GABAergic neuronal fates in cerebellum. *Neuron* 47, 201–213. doi: 10.1016/j.neuron.2005.06.007

Kani, S., Bae, Y. K., Shimizu, T., Tanabe, K., Satou, C., Parsons, M. J., et al. (2010). Proneural gene-linked neurogenesis in zebrafish cerebellum. *Dev. Biol.* 343, 1–17. doi: 10.1016/j.ydbio.2010.03.024

Katahira, T., Sato, T., Sugiyama, S., Okafuji, T., Araki, I., Funahashi, J., et al. (2000). Interaction between Otx2 and Gbx2 defines the organizing center for the optic tectum. *Mech. Dev.* 91, 43–52. doi: 10.1016/s0925-4773(99)00262-2

Kuraku, S., and Kuratani, S. (2006). Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zoolog. Sci.* 23, 1053–1064. doi: 10.2108/zsj.23.1053

Kuraku, S., Zmasek, C. M., Nishimura, O., and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* 41, W22–W28. doi: 10.1093/nar/gkt389

Lannoo, M. J., and Hawkes, R. (1997). A search for primitive Purkinje cells: zebrin II expression in sea lampreys (*Petromyzon marinus*). *Neurosci. Lett.* 237, 53–55. doi: 10.1016/s0304-3940(97)00802-1

Larsell, O. (1947). The cerebellum of myxinoids and petromyzonts including developmental stages in the lampreys. *J. Comp. Neurol.* 86, 395–445.

Leto, K., Arancillo, M., Becker, E. B., Buffo, A., Chiang, C., Ding, B., et al. (2016). Consensus paper: cerebellar development. *Cerebellum* 15, 789–828. doi: 10.1007/s12311-015-0724-2

Martí, E., and Bovolenta, P. (2002). Sonic hedgehog in CNS development: one signal, multiple outputs. *Trends Neurosci.* 25, 89–96. doi: 10.1016/s0166-2236(02)02062-3

McMahon, A. P., and Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* 62, 1073–1085. doi: 10.1016/0092-8674(90)90385-r

Murakami, Y., Ogasawara, M., Sugahara, F., Hirano, S., Satoh, N., and Kuratani, S. (2001). Identification and expression of the lamprey Pax6 gene: evolutionary origin of the segmented brain of vertebrates. *Development* 128, 3521–3531.

Nieuwenhuys, R. (1967). Comparative anatomy of the cerebellum. *Prog. Brain Res.* 25, 1–93. doi: 10.1016/S0079-6123(08)60962-0

Nieuwenhuys, R., Donkelaar, H. J. T., and Nicholson, C. (1998). *The Central Nervous System of Vertebrates*. Berlin: Springer.

Oisi, Y., Kakitani, O., Kuratani, S., and Ota, K. G. (2015). "Analysis of embryonic gene expression patterns in the hagfish," in *Situ Hybridization Methods. Neuromethods*, Vol. 99, ed. G. Hauptmann (New York, NY: Humana Press), 249–262.

Oisi, Y., Ota, K. G., Kuraku, S., Fujimoto, S., and Kuratani, S. (2013). Craniofacial development of hagfishes and the evolution of vertebrates. *Nature* 493, 175–180. doi: 10.1038/nature11794

Ota, K. G., Kuraku, S., and Kuratani, S. (2007). Hagfish embryology with reference to the evolution of the neural crest. *Nature* 446, 672–675. doi: 10.1038/nature05633

Pani, A. M., Mullarkey, E. E., Aronowicz, J., Assimacopoulos, S., Grove, E. A., and Lowe, C. J. (2012). Ancient deuterostome origins of vertebrate brain signalling centres. *Nature* 483, 289–294. doi: 10.1038/Nature10838

Pascual-Anaya, J., Sato, I., Sugahara, F., Higuchi, S., Paps, J., Ren, Y., et al. (2018). Hagfish and lamprey Hox genes reveal conservation of temporal colinearity in vertebrates. *Nat. Ecol. Evol.* 2, 859–866. doi: 10.1038/s41559-018-0526-2

Ravi, V., Bhatia, S., Shingate, P., Tay, B. H., Venkatesh, B., and Kleinjan, D. A. (2019). Lampreys, the jawless vertebrates, contain three Pax6 genes with distinct expression in eye, brain and pancreas. *Sci. Rep.* 9:19559. doi: 10.1038/s41598-019-56085-8

Reifers, F., Bohli, H., Walsh, E. C., Crossley, P. H., Stainier, D. Y., and Brand, M. (1998). Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis. *Development* 125, 2381–2395.

Sato, T., Araki, I., and Nakamura, H. (2001). Inductive signal and tissue responsiveness defining the tectum and the cerebellum. *Development* 128, 2461–2469.

Seto, Y., Nakatani, T., Masuyama, N., Taya, S., Kumai, M., Minaki, Y., et al. (2014). Temporal identity transition from Purkinje cell progenitors to GABAergic interneuron progenitors in the cerebellum. *Nat. Commun.* 5:3337. doi: 10.1038/ncomms4337

Shigetani, Y., Sugahara, F., Kawakami, Y., Murakami, Y., Hirano, S., and Kuratani, S. (2002). Heterotopic shift of epithelial-mesenchymal interactions in vertebrate jaw evolution. *Science* 296, 1316–1319. doi: 10.1126/science.1068310

Striedter, G. F., and Northcutt, R. G. (2020). *Brains Through Time: A Natural History of Vertebrates*. Oxford: Oxford University Press.

Sugahara, F., Aota, S., Kuraku, S., Murakami, Y., Takio-Ogawa, Y., Hirano, S., et al. (2011). Involvement of Hedgehog and FGF signalling in the lamprey telencephalon: evolution of regionalization and dorsoventral patterning of the vertebrate forebrain. *Development* 138, 1217–1226. doi: 10.1242/dev.059360

Sugahara, F., Murakami, Y., and Kuratani, S. (2015). "Gene expression analysis of lamprey embryos," in *In Situ Hybridization Methods. Neuromethods*, Vol. 99, ed. G. Hauptmann (Cham: Springer), 263–278.

Sugahara, F., Pascual-Anaya, J., Oisi, Y., Kuraku, S., Aota, S., Adachi, N., et al. (2016). Evidence from cyclostomes for complex regionalization of the ancestral vertebrate brain. *Nature* 531, 97–100. doi: 10.1038/nature16518

Sugahara, F., Murakami, Y., Pascual-Anaya, J., and Kuratani, S. (2017). Reconstructing the ancestral vertebrate brain. *Dev. Growth Differ* 59, 163–174. doi: 10.1111/dgd.12347

Tahara, Y. (1988). Normal stages of development in the lamprey, lampetra-reissneri (Dybowski). *Zoolog. Sci.* 5, 109–118.

Takio, Y., Kuraku, S., Murakami, Y., Pasqualetti, M., Rijli, F. M., Narita, Y., et al. (2007). Hox gene expression patterns in *Lethenteron japonicum* embryos–insights into the evolution of the vertebrate Hox code. *Dev. Biol.* 308, 606–620. doi: 10.1016/j.ydbio.2007.05.009

Takio, Y., Pasqualetti, M., Kuraku, S., Hirano, S., Rijli, F. M., and Kuratani, S. (2004). Evolutionary biology: lamprey Hox genes and the evolution of jaws. *Nature* 429:1 p following 262. doi: 10.1038/nature02616

Ueki, T., Kuratani, S., Hirano, S., and Aizawa, S. (1998). Otx cognates in a lamprey, *Lampetra japonica*. *Dev. Genes Evol.* 208, 223–228. doi: 10.1007/s004270050176

Wingate, R. J. (2001). The rhombic lip and early cerebellar development. *Curr. Opin. Neurobiol.* 11, 82–88. doi: 10.1016/s0959-4388(00)00177-x

Wullimann, M. F., Mueller, T., Distel, M., Babaryka, A., Grothe, B., and Köster, R. W. (2011). The long adventurous journey of rhombic lip cells in jawed vertebrates: a comparative developmental analysis. *Front. Neuroanat.* 5:27. doi: 10.3389/fnana.2011.00027

Yamada, M., Terao, M., Terashima, T., Fujiyama, T., Kawaguchi, Y., Nabeshima, Y., et al. (2007). Origin of climbing fiber neurons and their developmental dependence on Ptf1a. *J. Neurosci.* 27, 10924–10934. doi: 10.1523/jneurosci.1423-07.2007

Yu, T., Meiners, L. C., Danielsen, K., Wong, M. T., Bowler, T., Reinberg, D., et al. (2013). Deregulated FGF and homeotic gene expression underlies cerebellar vermis hypoplasia in CHARGE syndrome. *Elife* 2:e01305. doi: 10.7554/eLife.01305

Yu, T., Yaguchi, Y., Echevarria, D., Martinez, S., and Basson, M. A. (2011). Sprouty genes prevent excessive FGF signalling in multiple cell types throughout development of the cerebellum. *Development* 138, 2957–2968. doi: 10.1242/dev.063784

Zhang, H., Ravi, V., Tay, B. H., Tohari, S., Pillai, N. E., Prasad, A., et al. (2017). Lampreys, the jawless vertebrates, contain only two ParaHox gene clusters. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9146–9151. doi: 10.1073/pnas.1704457114

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Sugahara, Pascual-Anaya, Kuraku, Kuratani and Murakami. 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) and the copyright owner(s) 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 Cell and Developmental Biology

REVIEW
published: 06 September 2021
doi: 10.3389/fcell.2021.701779

Brain Sensory Organs of the Ascidian *Ciona robusta*: Structure, Function and Developmental Mechanisms

Paola Olivo‡, Antonio Palladino†‡, Filomena Ristoratore*‡ and Antonietta Spagnuolo*‡

Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Napoli, Italy

OPEN ACCESS

Edited by:
Stephanie Bertrand,
UMR 7232 Biologie Intégrative des Organismes Marins (BIOM), France

Reviewed by:
Takehiro Kusakabe,
Konan University, Japan
Jean-Philippe Chambon,
Université Paris-Sorbonne, France

***Correspondence:**
Filomena Ristoratore
filomena.ristoratore@szn.it
Antonietta Spagnuolo
antonietta.spagnuolo@szn.it

†Present address:
Antonio Palladino,
Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Napoli, Italy

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *28 April 2021*
Accepted: *12 August 2021*
Published: *06 September 2021*

Citation:
Olivo P, Palladino A, Ristoratore F and Spagnuolo A (2021) Brain Sensory Organs of the Ascidian Ciona robusta: Structure, Function and Developmental Mechanisms. Front. Cell Dev. Biol. 9:701779. doi: 10.3389/fcell.2021.701779

During evolution, new characters are designed by modifying pre-existing structures already present in ancient organisms. In this perspective, the Central Nervous System (CNS) of ascidian larva offers a good opportunity to analyze a complex phenomenon with a simplified approach. As sister group of vertebrates, ascidian tadpole larva exhibits a dorsal CNS, made up of only about 330 cells distributed into the anterior sensory brain vesicle (BV), connected to the motor ganglion (MG) and a caudal nerve cord (CNC) in the tail. Low number of cells does not mean, however, low complexity. The larval brain contains 177 neurons, for which a documented synaptic connectome is now available, and two pigmented organs, the otolith and the ocellus, controlling larval swimming behavior. The otolith is involved in gravity perception and the ocellus in light perception. Here, we specifically review the studies focused on the development of the building blocks of ascidians pigmented sensory organs, namely pigment cells and photoreceptor cells. We focus on what it is known, up to now, on the molecular bases of specification and differentiation of both lineages, on the function of these organs after larval hatching during pre-settlement period, and on the most cutting-edge technologies, like single cell RNAseq and genome editing CRISPR/CAS9, that, adapted and applied to *Ciona* embryos, are increasingly enhancing the tractability of *Ciona* for developmental studies, including pigmented organs formation.

Keywords: evolution, ascidians, pigmented cells, photoreceptor cells, molgula

INTRODUCTION

The sea squirt *Ciona intestinalis* represents a suitable model system for studying the transcriptional regulatory mechanisms exploited during chordate development (Satoh et al., 2003), due to its phylogenetic position (belonging to chordates as sister group of vertebrates) and to a number of peculiar features. These include the rapid embryogenesis as well as the small size of its larval stage (∼2,500 cells when fully developed) showing only six tissue types, as notochord, muscle, epidermis, endoderm, mesenchyme and nervous system together with a tunic, made of cellulose like material that covers the entire body of both, larva and adult, hence the name tunicates. *Ciona* is suitable for genetic studies since its small and compact genome (160 Mb and contains ∼15,500 genes) (Dehal et al., 2002) often contains single copy of genes present in multiple copies in the genome of vertebrates. This characteristic permits to reveal the function of the genes involved in fundamental steps of the developmental programs in the lineage of chordates. Furthermore, the simplicity of *Ciona* larval structures, in particular of its Central Nervous System (CNS), permits to depict the genetic programs adopted by a single blastomere in order to build up specific structures, tissues, organs present at the larval stage. This simplicity is undoubtedly a huge advantage compared to the complexity of vertebrates. In this context, the ascidian pigmented sensory organs represent, for more than two decades, a fertile ground for deep studies aimed at revealing their structure and function and deciphering the molecular mechanism leading to their development since the gastrula stage.

Inside ascidian group, recently, two cryptic species of *Ciona intestinalis*, types A and B (Caputi et al., 2007; Sato et al., 2012; Roux et al., 2013; Bouchemousse et al., 2016) have been described and proposed to be renamed, based on a taxonomic study (Brunetti et al., 2015), as *Ciona robusta* (*Ciona intestinalis* type A) and *Ciona intestinalis* (*Ciona intestinalis* type B). For the sake of simplicity and to avoid confusion, in this review the general term *Ciona* will be used for both species.

STRUCTURE AND FUNCTIONS OF ASCIDIAN PIGMENTED SENSORY ORGANS

In the tadpole larva of *Ciona*, the CNS is formed, from the anterior to the posterior, by three structures: the brain vesicle (or sensory vesicle), the visceral ganglion and the caudal nerve cord (Katz, 1983; Nicol and Meinertzhagen, 1991; Meinertzhagen et al., 2004).

The ascidian larval brain (or sensory vesicle), located in the larval trunk, contains almost two thirds of the total number of the cells constituting the whole CNS (215/335). The two most evident structures present in the larval brain are the pigmented sensory organs, the otolith and the ocellus (**Figure 1**). The most ventral one is the otolith, made of a pigmented cell, two associated ciliated cells and two glutamatergic antennae sensory neurons. The function of the two otolith-associated ciliated cells is not known, but they could be involved in detecting movements of the otolith cell. The otolith cell is anchored to the cavity of the brain vesicle by a stalk and can move inside the cavity; furthermore, the otolith cell can change its density since its melanin granules are able to chelate metal ions. The otolith is responsible for gravity perception and pigmentation is essential for proper geotactic behavior (Tsuda et al., 2003; Sakurai et al., 2004; Jiang et al., 2005). The two antenna cells are able to detect gravity and project their axons to eleven interneurons located in the posterior brain that in turn is connected, through the neck, to the motor ganglion (Ryan et al., 2016). Ascidian larvae exhibit conserved phototactic and geotactic behaviors. Free-swimming larvae show geonegative response for most of the larval dispersal period until shortly before settlement, when they start to swim toward gravity to attach on a substrate and start metamorphosis. It has been demonstrated that pigmented cells of the ocellus and otolith are necessary for sensing light and gravity, respectively, so guiding the swimming behaviour before settlement (Tsuda et al., 2003; Sakurai et al., 2004).

The ocellus is a multicellular structure necessary for light perception and is located in the sensory vesicle on the dorsal-right side. Ocellus is made of a cup-shaped pigment cell, 3 lens cells (that are not homologous of vertebrates lens) and is associated with almost 30 photoreceptor cells, gathered into 3 groups, based on typical morphological characteristics: Group I (PR-I), Group II (PR-II), and Group III (PR-III) photoreceptor cells (Horie et al., 2008; **Figure 1**). The Group I, made of 18–23 cells, and the Group II, made of 8–11 cells, are both in proximity of the ocellus pigment cell. The Group I is in close contact with the pigment cell and the outer segment of each cell is arranged in rows inside the pigment cup. Their function is to perceive the light directed to the cup through the three lens cells, while the pigment cell protects them from light coming from other directions. The Group II photoreceptors are in more anterior position and outside of the pigment cup, without being protected by the pigment cell, thus sensing the light from any direction. These morphological differences and location did suggest that Groups I and II have distinct functions (see the paragraph Visuomotor and gravitaxis circuits). The Group III consists of 6–7 photoreceptor cells located far from the ocellus pigment cell and closer to the otolith. The shape of these photoreceptor cells is circular, and their outer segments are exposed into the lumen of the sensory vesicle. The Group III alone is not sufficient to evoke a photoresponse behavior in *Ciona* larvae, as demonstrated by laser ablation experiments of Groups I and II photoreceptor cells (Horie et al., 2008). This, however, does not exclude the contribution of Group III to the process. Group III photoreceptor cells differentiate later, compared to Group I and II, as revealed by immunostaining experiments with anti-Arrestin and anti-Opsin1 antibodies. The staining with both antibodies persists during early stages of metamorphosis, thus suggesting the potential function of Group III photoreceptor cells in the late larval-early metamorphosis stages. Close to the Group III is present a cluster of cells, called coronet cells, able to produce dopamine (DA), whose function is still unknown (**Figure 1B**; Eakin and Kuda, 1971; Nicol and Meinertzhagen, 1991; Moret et al., 2005). Some authors (Eakin and Kuda, 1971) suggested that these cells could be involved in pressure detection (hence the name of pressure organ), while others (Razy-Krajka et al., 2012) have speculated their implication in the photic response; however, in both cases these hypotheses need to be supported by experimental data.

It is intriguing to note that *Ciona* ocellus and vertebrate eyes share specific features of photo-transduction, which led to suppose a common origin of both structures (Kusakabe et al., 2001; Sato and Yamamoto, 2001). In particular, *Ciona* photoreceptors are hyperpolarizing and ciliary, like those of vertebrates (Gorman et al., 1971) and use visual opsins, G-protein coupled receptors, and visual arrestins, small proteins needed to regulate opsin signal transduction, during photo-transduction process, as their vertebrate counterpart (Arshavsky et al., 2002; Blomhoff and Blomhoff, 2006). Notably, both genes, precisely *Opsin1* (three *Opsin* genes are present in *Ciona* genome) and *Arrestin* (one *Arrestin* gene is present in *Ciona* genome), are expressed in the three groups of *Ciona* photoreceptor cells (Kusakabe et al., 2001; Horie et al., 2002, 2008; Nakashima et al., 2003).

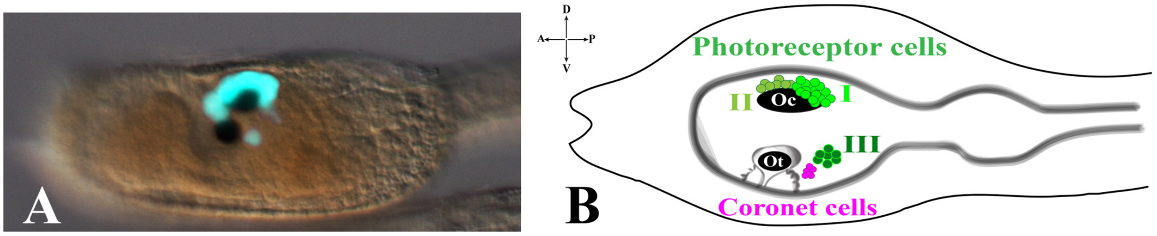


FIGURE 1 | Pigmented organs in the sensory vesicle. **(A)** pArr>eGFP expression in photoreceptor territories of *Ciona* larval trunk (Yoshida et al., 2004) (photo from our lab). **(B)** Schematic representation of larval trunk. Ocellus (Oc) and otolith (Ot) pigment cells. Photoreceptor cell types I, II are associated with ocellus pigment cell, while photoreceptor cell type III is in close proximity to otolith pigment cell and coronet cells.

PHOTORECEPTOR CELL LINEAGE IN *CIONA*

In *Ciona*, many developmental processes have been defined thanks to the precise outline of the cell lineages of most tissues and organs. In this regard, the most striking example is represented by the CNS precursors that, since the late gastrula stage (neural plate), are structured like a grid, formed by eight columns (two bilaterally symmetrical four columns) and six rows (I-VI) (**Figure 2**). In particular, the posterior rows I and II, formed by the anterior vegetal A-line lineage blastomeres, contains the precursors of the caudal nerve cord, motor ganglion and posterior sensory vesicle of the larvae (Imai et al., 2009). The rows III-VI, which include the anterior ectodermal a-line lineage, will form the anterior nervous system. In detail, the anterior sensory vesicle is formed by descendants of row III a-line blastomeres, cells of row IV contribute to the stomodeum as well as the anterior brain, while the progeny of rows V and VI blastomeres contributes to the palps as well as to the epidermal sensory neurons in the trunk.

Concerning the photoreceptor cells, their precursors have long been identified in the right a9.33 and a9.37 blastomeres (**Figure 2**, green and red, respectively), while the precursors of the coronet cells were first thought to be the left a9.33 and a9.37 blastomeres (**Figure 2**, light green and pink, respectively) (Nicol and Meinertzhagen, 1991; Cole and Meinertzhagen, 2004).

However, the whole scenario, more recently, has been completely changed (Gainous et al., 2015; Oonuma et al., 2016), and the precursors of photoreceptor cells have been recognized to be in more posterior regions of the neural plate, the medial regions of row II. All this was achieved by labeling specific blastomeres of "non-dechorionated" embryos to trace the developmental fates of each neural plate cell, from the late gastrula up to the larval stage. In these experiments, the authors used Kaede photoactivatable reporter downstream from the regulatory regions of *Dmrt1* gene, specific for a-lineage (Wagner and Levine, 2012), and *FoxB* gene (Imai et al., 2004), specific for A-lineage, since the 32–64 cell stage. Kaede photoconverted larvae were then analyzed by co-immunostaining with anti-Arrestin and anti-Opsin1 antibodies. In these experiments, the chorion was not removed in order to avoid any aberrant left-right asymmetry (Shimeld and Levin, 2006), often detected in "dechorionated" embryos, which in turn could alter the tracing of the correct cell lineage. The results of this study indicated that Group I, Group II and Group III photoreceptor cells originate from the A-lineage blastomeres and, specifically, those of Group I and Group II are the descendants of the right A9.14 (**Figure 2**, blue) cell and those of Group III derive from the right A9.16 cell (**Figure 2**, yellow).

VISUOMOTOR AND GRAVITAXIS CIRCUITS

Recently, the visuomotor responses of *Ciona* larva to rapid light dimming (or shadow response) and negative phototaxis

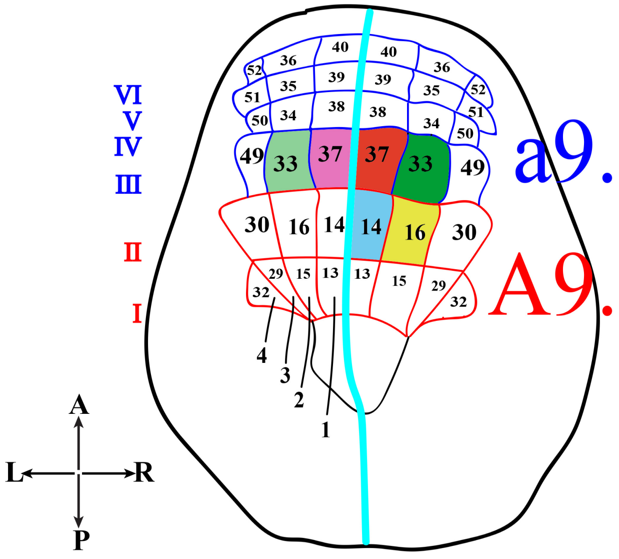


FIGURE 2 | Schematic representation of *Ciona* neural plate stage. Neural plate, dorsal view, with 4 columns (1–4) and six rows (I-VI). The turquoise line separates the two bilaterally symmetrical blastomere pairs. The former a9-lineage photoreceptor cell precursors, right a9.33 and a9.37, are in green and red, respectively, while the supposed precursors of coronet cells, left a9.33 and a9.37, are in light green and light pink. The current A9-lineage Group I and II (A9.14) photoreceptor cell precursors is indicated in blue and Group III (A9.16) photoreceptor cell precursors in yellow.

have been the subject of a detailed study (Salas et al., 2018). The authors demonstrated that *Ciona* larvae tend to escape from the light and exhibit a negative phototaxis, which depends on the intensity of the illumination and is characterized by sustained directional swims. That is: lower illumination/reduced negative phototaxis; higher illumination/increased negative phototaxis. The negative phototaxis is "age-dependent": young larvae (20.5 hpf at 18°C) lacked this behavior but were able to respond to dimming light; older larvae (from 23 hpf at 18°C) showed instead both behaviors (phototaxis and shadow response) that resulted to be quantitatively and qualitatively different. Indeed, the swims evoked by light dimming were more tortuous compared to the swims evoked by directional light (negative phototaxis) and the dimming response appeared independent of light direction. Interestingly, the albino mutant pristine (*prs*), which does not contain pigmented cells in the brain vesicle but has photoreceptors, showed no phototactic behavior, supporting previous assertions that phototaxis is linked to Group I photoreceptor cells associated to a pigmented cell (Mast, 1921; Svane and Young, 1989; Jiang et al., 2005). *Prs* larvae, however, showed an elevated shadow response and a lower tortuosity of the resulting swims, compared to control larvae, thus indicating their positive photoresponse.

Ciona connectome, recently mapped, revealed the smallest CNS known in any chordate, with only 177 neurons, distributed in the three structures (Ryan et al., 2016). These neurons can be split into at least 25 types and each of them has, on average, 49 synapses with other cells. Thus, despite the small number of neurons, the neuron network shows a substantial complexity. These minimal and specific neural circuits, simple versions of the intricate networks of vertebrates, offer a unique opportunity to describe basal features of chordate CNS and to study the influence of each single neuron on the behavior of *Ciona* larva.

In this regard, a fine study depicted the minimal circuit in which the activation of motor neurons is linked to both the Group I and Group II photoreceptor cells (**Figure 3**; Ryan et al., 2016). Both Groups extend synapses in the posterior Brain Vesicle (BV). Group I (PR-I) transmits inputs to photoreceptor Relay Neurons (prRN) which in turn connects to the paired right/left Motor Ganglion (MGIN) interneurons and then to motor neurons (MN, five on each side). Group I (PR-I), together with Group II (PR-II), connects also to the photoreceptor Ascending MG RNs (pr-AMG RN) which are so-named because they, unlike the prRNs, receive input from the Ascending MG peripheral interneurons (AMG neurons; not shown in **Figure 3**). Notably, the pr-AMG RNs and the prRNs are also highly interconnected. Like the prRNs, the pr-AMG RNs extend synapses to the left and right Motor Ganglion InterNeurons (MGIN) which in turn are connected to the paired right and left motor neurons (MN). Thus, a complete visuomotor circuit, moving from photoreceptors up to the muscle target cells, is present in *Ciona* connectome (**Figure 3**).

This simple circuit has been recently connected with neurotransmitter use and behavioral observations, thanks to a study by Kourakis and his collaborators (Kourakis et al., 2019). The authors, by fine and elegant experiments, inferred a model in which *Ciona* exploits the PR-Is to sense the direction of light and PR-IIs to perceive the changes in ambient light. The PR-I circuit uses glutamate, is excitatory and projects, as previously reported, to both cholinergic prRNs (expressing the glutamate AMPA receptor (AMPAR)) (Okamura et al., 2005) and GABAergic pr-AMG RNs (**Figure 3**), which in turn synapses onto the cholinergic MGINs then connected to MNs. Notably, the interconnection between pr-AMG RNs and prRNs suggests that these clusters are not simple conveyors of information, from photoreceptor cells to motor neurons, but they can play a primary role in visual processing.

The PR-II circuit uses mainly GABA and synapses onto the GABAergic pr-AMG RNs, thus indicating that the PR-II output to the pr-AMG RNs is predominantly inhibitory (**Figure 3**). Notably, some PR-IIs co-express VGAT and VGLUT and this, although speculative, could be related to the need of fine tuning excitatory/inhibitory balance.

The light response is further integrated with the gravity response to generate a complex behavior that permits the rapid reorientation of larvae in response to dimming. Gravitaxis in *Ciona* is controlled by the sensory otolith and the minimal gravitaxis circuit, leading to motor neurons, is made up of the

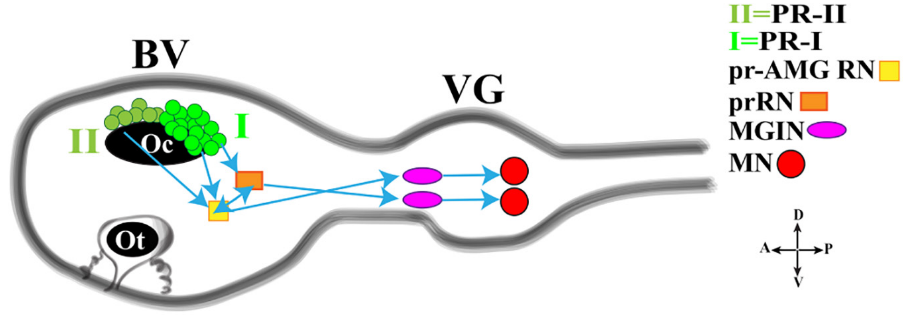


FIGURE 3 | Cartoon of a *Ciona* tadpole larva Central Nervous System showing the minimal visuomotor circuit. D., dorsal; V., ventral; A., anterior; P., posterior; PR-II, photoreceptor group II; PR-I, photoreceptor group I; pr-AMG RN, Photoreceptor Ascending Motor Ganglion Relay Neuron; prRN, photoreceptor Relay Neuron; MGIN, Motor Ganglion InterNeuron; MN, Motor Neuron. BV, Brain Vesicle; VG, Visceral Ganglion. Adapted from Kourakis et al. (2019).

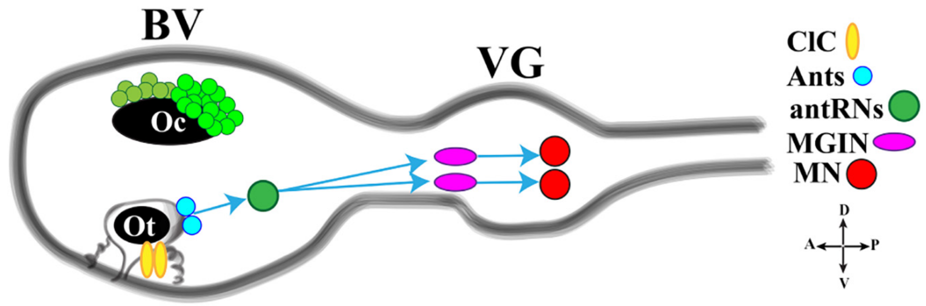


FIGURE 4 | Cartoon of a *Ciona* tadpole larva Central Nervous System showing the minimal gravitaxis circuit. D., dorsal; V., ventral; A., anterior; P., posterior; CIC, ciliated cells; Ants: glutamatergic antennae sensory neurons; antRNS, GABAergic antenna Relay Neurons; MGIN, Motor Ganglion Inter Neuron; MN, Motor Neuron. BV, Brain Vesicle; VG, Visceral Ganglion. Adapted from Kourakis et al. (2019) and Bostwick et al. (2020).

otolith cell connected to the two glutamatergic antenna cells (Ants) which synapse onto the GABAergic antenna relay neurons (antRNs) asymmetrically connected to the right cholinergic MGINs and in turn to MNs (**Figure 4**).

Recently, Bostwick et al. (2020) investigated the neural circuits of gravitaxis in *Ciona* larvae using connectomic and neurotransmitter data, coupled with sophisticated behavioral assays. To sum up, the authors inferred that the gravitaxis circuit induces curved or tumbling (asymmetric) swimming when the light is off and the larva is head down or sideways, while, when the larva is facing upward, this circuit is almost inactive. The gravitaxis appears actively repressed until light off and the repression state is supported by the intersection between gravitaxis and photoreceptor Group II. As previously reported, the PR-II output to the pr-AMG RNs is predominantly inhibitory. Interestingly, pr-AMG RNs synaptically contact the antRNs, which are targets of the gravitaxis circuit in the posterior BV. These synaptic contacts thus result in a plausible circuit model in which inhibitory input from PR-II suppresses gravitaxis until dimming. Once the light is off, the inhibition is released, on both MGINs and antRNs, and larva initiates the asymmetric swimming. The gravitaxis circuit most probably is involved in modulating swims, by inducing curved and tumbling movements thanks to the asymmetric connections to motoneurons. Notably, the gravitaxis circuit seems to maturate during larval growth between 21 and 25 hpf.

Collectively these studies demonstrate the power of genetics, combined with behavioral analyses and connectomic data, to track down relevant functional information.

MOLECULAR MECHANISMS CONTROLLING PHOTORECEPTORS DIFFERENTIATION

Rx (*retinal homeobox*)/*Rax* (retina and anterior neural fold homeobox) genes play pivotal roles in eye formation in vertebrates, since knockout of *Rx* genes results in the absence of eyes (Bailey et al., 2004). At molecular levels knockdown of *Rx* causes a reduction in the expression of the genes encoding *Arrestin* and *Rhodopsin* (Pan et al., 2010). A study on *Ciona* indicated that the gene orthologous to *Rx* exerts almost the same functions as its vertebrate counterpart. Loss of function experiments showed, indeed, that larvae lacking *Rx* do not develop photoreceptor cells, as revealed by reduced or absent expression of *Arrestin* and *Opsin* genes, and are unable to respond to light stimuli variations (D'Aniello et al., 2006). However, this study suggested that *Rx* is expressed in a-lineage cells at the tailbud stage and not in the A-lineage from which the photoreceptor cells develop, as recently assessed. To resolve this issue, the expression profile of *Rx* in *Ciona* has been recently reanalyzed, with the aim to reveal the possible presence of *Rx* in the A-lineage and substantiate its role in photoreceptor cells formation (Oonuma and Kusakabe, 2019). The authors confirmed that the expression of *Rx* starts in the a-lineage at the early tailbud stage, as previously demonstrated, but they revealed that, from the middle-late tailbud stage (13th generation cells), *Rx* expression starts in the most anterior A-line blastomeres derived from the right A9.14 cell. During next cell divisions, the *Rx* staining expands posteriorly to the whole descendants of the right A9.14 cells, while continuing to be expressed, anteriorly, in the right a-lineage descendants. Thus, the detection of *Rx* in the A9.14 lineage fits now with its involvement in photoreceptor cells differentiation.

Concerning *Rx* regulation, the authors speculated that Homeodomain and Sox Binding Sites are important for the activation of *Ciona Rx* in photoreceptor lineage, while *Onecut*, previously demonstrated to be involved in *Rx* expression (D'Aniello et al., 2011), may regulate *Rx* in the a-lineage cells, but not in the A-lineage cells for two main reasons. *Onecut* expression has not been detected so far in the A-lineage photoreceptor progenitor cells and mutations affecting the Homeodomain Binding sites (BSs), but not Onecut BSs, in the proximal enhancer region, significantly reduced transgene activity in the A-lineage cells. Thus, it appears that a combined action of activators and repressors is responsible for proper expression of *Rx* in photoreceptor lineage. Homeodomain transcription factors seem to be required for *Rx* activation, while members of *Sox* family could be involved in the repression of ectopic *Rx* expression. However, further studies are required to verify the effective role of these candidate genes in *Rx* regulation and to elucidate the detailed mechanism, acting upstream from *Rx*, controlling early specification of photoreceptor cell lineages.

GENE REGULATORY NETWORK UNDERLYING PIGMENTED CELLS SPECIFICATION

In both *Ciona* and *Halocynthia roretzi*, cell lineage of pigmented cells becomes directed toward this fate, starting from gastrula stage, in two bilateral symmetric blastomeres, the a8.25 pairs (Nishida, 1987; Cole and Meinertzhagen, 2004). Then, the a8.25 cells divide once giving rise to the a9.49 and a9.50 cell pairs and, as the development proceeds, the a9.50s (positioned in the Row IV of the neural plate) will give rise to the anterior part of the CNS, while the a9.49s (located in the Row III) will retain the fate of pigment cells (**Figure 5A**). Both a9.49s express *Tyrosinase (Tyr)* and *Tyrosinase related protein genes* (*Tyrp*) that are essential for melanogenesis (Tief et al., 1996; Caracciolo et al., 1997; Esposito et al., 2012; Haupaix et al., 2014; Racioppi et al., 2014). A further division of the a9.49s cell pairs, at the mid-neurula stage, give rise to the a10.97s and a10.98s pairs, which after a further division, at the early tailbud stage, will form eight post-mitotic cells (Haupaix et al., 2014; Racioppi et al., 2014) that, according to cell nomenclature described in Cole and Meinertzhagen (Cole and Meinertzhagen, 2004), are named a11.193/a11.194 and a11.195/a11.196 pairs, respectively. The eight Pigment Cell Precursor (PCP) migrate and intercalate, during neural tube closure, aligning in a single row, along the anterior-posterior axis. Among them, the fate of pigmented cells remains restricted to the a11.193 pair and the final choice among ocellus or otolith pigment cell fate is based on the anterior-posterior position of these cells. The most anterior cell migrates inside the sensory vesicle and develops as otolith pigmented cell, while the posterior one will form the ocellus pigment cell (Nishida and Satoh, 1989; Darras and Nishida, 2001).

Pigment cell precursor formation represents one of the first example of inducible cell fate in *Ciona*. It has been demonstrated that FGF signal is one of the actors responsible for this induction. Indeed, at the gastrula stage, the a9.49 cells, in row III, are both located in a position close to blastomeres of row II expressing the signaling factors Ci-fgf8/17/18 and Ci-fgf9/16/20. Moreover, the known downstream effector of FGF signaling, the transcription factor Ci-ERK1/2, is present in its activated form (phosphorylated) in the cells of the neural plate row III, including the a9.49s pigment cell precursors (Hudson et al., 2007; Racioppi et al., 2014).

The requirement of the FGF signaling for proper specification of the pigment cells in *Ciona* sensory organs has been demonstrated in several elegant experiments (Hudson et al., 2007; Squarzoni et al., 2011; Racioppi et al., 2014). FGF signaling, via the MAPK/ERK cascade and the activation of Ets1/2 transcription factor, is required for pigment cell specification in *Ciona*. Ets1/2 transcripts are present in the a9.49 cells and are inherited by their descendants, a10.97 and a10.98 blastomere pairs. Intriguingly, morphogenetic rearrangements put the a10.98 blastomeres more distant from FGF sources, in a rostral position (**Figure 5B**), while the a10.97s remains in close proximity of row II cells producing FGF secreted factors. As it is well known, FGF molecules are able to extend their function at short range from the target cells (Tassy et al., 2006), hence Ets1/2 factor results activated (phosphorylated) solely in the a10.97 blastomeres and able to initiate the transcription of the genes necessary for pigment cells specification. Thus, the key event in the cascade is just the activation, through phosphorylation, of Ets1/2. Differential activation of ERK1/2 in the a9.49 cell pairs is related to MAPK/ERK phosphorylation cascade and finely regulated by the presence of the Eph/ephrin signal, which has been shown to act, during gastrulation, by limiting the activation of ERK1/2 factor into the a10.97 cell pairs (Hudson et al., 2007). Indeed, interference with the Eph/ephrin signal results in extranumerary pigment cells at the larval stage and ectopic expression of *Trp* in a10.98 cells (Haupaix et al., 2014). Collectively, these data point to a fundamental role played by FGF/MAPK/ERK/Ets signaling in pigment cells specification.

Another relevant factor involved in pigment cells formation in *Ciona* is Tcf, a downstream effector of Wnt signaling (Molenaar et al., 1996). Notably, Ciona *Tcf* is transcriptionally regulated by FGF signaling, through direct control of Ets1/2 factor, and this interaction makes the a9.49 pairs "competent" to respond to Wnt signal (Squarzoni et al., 2011). The crosstalk of two signaling cascades at the transcriptional level represent a new way to study how different inputs can intertwine to control the specificity of cellular responses.

Further data on pigment cell differentiation have been obtained employing lineage-specific transcription profiling of PCP upon interference with FGF signaling. Our group demonstrated that FGF signaling induces pigment cell identity in the blastomeres that, in absence of this signal, would adopt an anterior neural fate (Racioppi et al., 2014). Among the genes identified in this analysis, *Rab32/38*, a small GTPase involved in molecular trafficking, has been studied and demonstrated, by mutational analyses, to play a role in the pigmentation of *Ciona* larval sensory organs (Racioppi et al., 2014, 2019b).

In *H. roretzi,* a further cascade involved in the pigmentation program is the BMP (Bone Morphogenetic Protein) signaling. The authors demonstrated that one or no pigment cells are formed in the larval sensory vesicle by overexpressing *BMPb* (Miya et al., 1997). When pigmented cells precursors are aligned along the dorsal neural tube, BMP signal, in antagonism with Chordin, that is present in the cell located just posterior to the PCPs, is responsible for the ocellus/otolith choice (Darras and Nishida, 2001).

Differently from *H. roretzi*, BMP/Chordin antagonism does not seem to be involved in pigment cells fate choice in *Ciona*. Data from Abitua and collaborators indicated a role by Wnt signaling in *Ciona* otolith vs. ocellus differentiation. The authors demonstrated that, at the tailbud stage, the ocellus precursor (a11.193) expresses *Tcf* gene (Abitua et al., 2012) and that *Wnt7* (one of the Wnt genes present in *Ciona* genome) is expressed, along the dorsal midline, in the blastomere just behind the ocellus precursor. If Wnt7 is misexpressed in more anterior position, both a11.193 PCP blastomeres give rise to ocelli. On

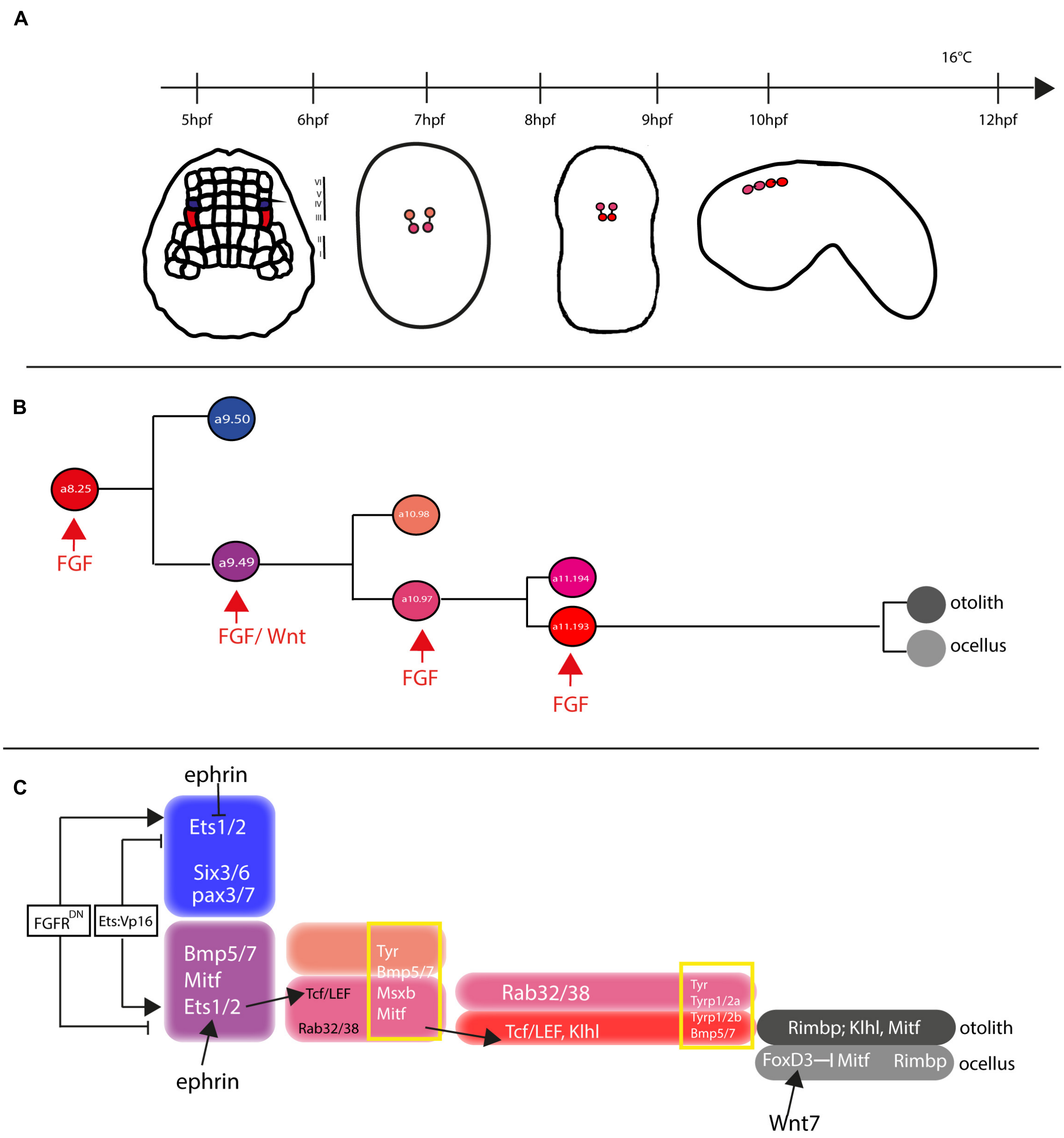

A
16°C
5hpf
6hpf
7hpf
8hpf
9hpf
10hpf
12hpf
VI
V
IV
III
II
I
B
a9.50
a8.25
FGF
a9.49
FGF/ Wnt
a10.98
a10.97
FGF
a11.194
a11.193
FGF
otolith
ocellus
C
ephrin
Ets1/2
Six3/6
pax3/7
FGFR^DN
Ets:Vp16
Bmp5/7
Mitf
Ets1/2
ephrin
Tcf/LEF
Rab32/38
Tyr
Bmp5/7
Msxb
Mitf
Rab32/38
Tcf/LEF, Klhl
Tyr
Tyrp1/2a
Tyrp1/2b
Bmp5/7
Rimbp; Klhl, Mitf
otolith
FoxD3—I Mitf
Rimbp
ocellus
Wnt7

FIGURE 5 | Otolith/ocellus cell fate in *Ciona robusta*. **(A)** Summary of pigmented cell lineage development within CNS, with the developmental stages (hours post fertilization, hpf) indicated the timeline [from Racioppi et al. (2014)]. **(B)** Fate determination of otolith and ocellus precursors from a9.49 blastomeres, under the control of FGF [adapted from Racioppi et al. (2014)]. **(C)** Reconstruction of genes involved in controlling otolith/ocellus determination. Adapted from Racioppi et al. (2014).

the other hand, interference with *Tcf* in PCP blastomeres induces differentiation of otolith in both 11.193s.

Furthermore, a suppressive role for the transcription factor *FoxD* has been identified. *FoxD*, which is specifically expressed in ocellus precursor at the tailbud stage, exerts its function by blocking the expression of *Mitf* transcription factor in this cell, thus attenuating the formation of melanin granules and leading to the formation of the less pigmented ocellus. On the other hand, *Mitf* expression, not attenuated by the presence of FoxD factor, leads to the sustained formation of melanin granules in the densely pigmented otolith.

Taken together, these data suggest a possible gene regulatory network (GNR) responsible for the differentiation of *Ciona* otolith and ocellus pigment cells (**Figure 5C**) in which *Wnt7* activates *FoxD* in the ocellus precursor, which in turn suppresses *Mitf* expression in this cell (Abitua et al., 2012; **Figure 5C**).

Recently, a new player in the specification of PCPs, *Klhl21/30*, has been identified. This gene, belonging to the *Kelch* family of genes, has been shown to be present in a dynamic pattern in *Ciona* PCPs, starting to be expressed in both the a10.49s and then becoming restricted to the otolith precursor at the tailbud stage. To place this gene inside the GRN of the PCPs, its regulatory region has been studied. The data demonstrated that *Klhl21/30* minimal key *cis*-regulatory element, able to drive its expression specifically in *Ciona* otolith, is controlled by Mitf, Msx, and Dmrt transcription factors, that work synergically to control the specific expression of the gene (Coppola et al., 2020).

SENSORY ORGANS IN MOLGULIDAE: A SPECIAL CASE

Despite the conservation in the general organization of the body plan of the swimming ascidian larvae, some species have developed an anural larval stage (Lacaze-Duthiers and de Lacaze-Duthiers, 1858; Berrill and Watson, 1931; Jeffery and Swalla, 1990).

About 20 species, among tunicates, present this type of larva and most of them belong to the Molgulidae family, where tail loss has occurred, independently, multiple times (Jeffery and Swalla, 1990; Huber et al., 2000). The reason why the molgula have developed a larval stage lacking some of the chordate distinctive characteristics is still not clear. Some authors suggested that the specific habitats, where flat sand is predominant and to which these species are adapted, could be at the basis of these changes, because in these natural conditions the dispersal phase is not useful and so dispensable (Huber et al., 2000).

Together with tail, anural larvae have lost most of the characteristics more directly associated with swimming, as the capability to form tail muscles (Whittaker, 1979; Swalla and Jeffery, 1990, 1992; Bates and Mallett, 1991; Bates, 1995; Tagawa et al., 1997) and this is linked to the inactivation, by pseudogenization, of genes like muscle actin, as demonstrated in *Molgula occulta* (Kusakabe et al., 1996; Jeffery et al., 1999).

Regarding sensory organs, it has been described that most of the molgulidae larvae present only the otolith, while the ocellus pigmented organ is absent, thus becoming unable to respond to light. In other species, like *Molgula occulta*, both pigmented organs are absent (Berrill and Watson, 1931) or vestigial (Swalla and Jeffery, 1992) and this loss has been associated with the absence of functional genes coding for *Tyrosinase* and *Tyrosinase related proteins* (*Tyrp*) (Racioppi et al., 2017). As an example, in *M. occulta* these genes show several mutations, like bases insertion or deletion, leading to the formation of premature stop codons into the mRNA, which becomes unable to code for functional enzymes (Racioppi et al., 2017).

These mutations are most probably still accumulating in *Molgula* genome, due to a loss of selective pressure on genes responsible for the pigmented organs formation (Racioppi et al., 2017).

NEW APPROACHES FOR THE STUDY OF ASCIDIANS SENSORY ORGANS

Developmental biology approaches take great advantages on the interference with a gene expression (silence or overexpression) to understand its function during development of a specific cell lineage in model organisms. During the past two decades a wide range of genetic tools have been developed and, at the same time, a series of new model organisms have established themselves as good models for the study of developmental mechanisms, physiology and behavior.

CRISPR/Cas

Most of the methods used to modify a specific sequence of DNA rely on the principle of site-specific DNA recognition by oligonucleotides, small molecules, or self-splicing introns, such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). Recently, CRISPR/Cas9 technique has been shown to function, with high fidelity and efficiency, in *Ciona* (Stolfi et al., 2014; Gandhi et al., 2017). This technique, exploited by the prokaryotic adaptive immune system, has been modified for the use in eukaryotic cells. CRISPR/cas9 technology is based on the use of an oligonucleotide, called short guide RNA (sgRNA), with a sequence of 20 nucleotides identical to the target DNA sequence, and an RNA mediated nuclease Cas9. The interaction of the sgRNA with its target genomic DNA sequence guides the Cas9 nuclease to this site. Cas9 unwind the DNA and cleave both strands. During the resulting repair mechanism, by homologous recombination or non-homologous end joining, insertions and/or deletions can accumulate at the target site, thus creating a mutation in the original sequence of the gene and its transcript (Jinek et al., 2012). In *Ciona* the embryos are transfected by electroporation method, allowing tissue-specific disruption of the gene of interest. Fertilized embryos are electroporated at one-cell stage with two plasmids, one driving the zygotic expression of Cas9 protein under a lineage specific promoter of choice, the other one with the selected sgRNAs under U6 ubiquitous promoter (Nishiyama and Fujiwara, 2008; Stolfi et al., 2014; Gandhi et al., 2017). The phenotypes are normally observed in F0 generation, allowing a very rapid analysis of the mutated larvae. The disadvantage of this technique in this model organism is related to the difficulties in obtaining transgenic strains of *Ciona* in order to study the mutations obtained by gene editing in the F1 generation. However, the phenotypes in F0 generation can be exploited as a powerful tool to investigate tissue-specific functions of a gene during development. In this regard, CRISPR/Cas9 technology has been used in *Ciona* to study some aspects of the Central and Peripheral nervous system development (Stolfi et al., 2014). As an example, targeted interference with FGF signaling by CRISPR/Cas9 has been instrumental to study some aspects of CNS development by using Bipolar Tail Neuron (BTNs) as an accessible model system for neurogenesis (Kim et al., 2020). Regarding cell fate determination of the otolith and ocellus in *Ciona,* preliminary results have been obtained studying, by CRISPR/Cas9 technique, *Mitf* gene, a key regulator of melanocyte development and melanoma in vertebrates (Levy et al., 2006). The study revealed that *Mitf* is a fundamental transcription factor involved in KLHL *21/30* (a specific marker of otolith) expression in *Ciona* (Coppola et al., 2020). Further studies, using transcriptomic approaches on *Mitf* mutated embryos, will be useful to study the gene regulatory network involved in pigmented cells differentiation.

scRNAseq

The simple organization and morphology of ascidian embryos highlights the crucial role of cell lineage during animal development. These characteristics, together with the advance in recent techniques, as single cell sequencing (scRNAseq), give the opportunity to investigate the transcriptome of a specific cell population. The use of scRNAseq technique is expanding in many model organisms including *Ciona*, where encouraging preliminary studies have been carried out. Extensive transcriptome trajectories, regulatory cascades and provisional gene networks for over 60 cell types, including pigmented cell lineage, were rebuilt by analyzing the transcriptome profiles of individual cells from gastrulation, 110 cell-stage, to larval stages (Cao et al., 2019). The use of scRNAseq has been applied to describe the transcriptome of *Ciona* different area of nervous system. In particular, for larval brain, 10 different clusters were identified and, within them, the most representative neural tissues were re-clustered, like dorsolateral brain, sensory/brain vesicle wall, epidermal neurons and ventral brain (Sharma et al., 2019). RNA sequencing allowed to explore the features of the lateral plate ectodermal of *Ciona*, which have common traits with neural plate ectoderm in vertebrates; both systems reveal similarities in the compartmentalization and regulatory program of *Six 1/2, Pax3/7* and *Msx* expression. These results support the hypothesis that the compartmentalization of lateral plate ectoderm preceded the origin of vertebrates (Horie R. et al., 2018). Moreover, the regulatory network underlying the specification of coronet cells, a class of dopaminergic neurons in *Ciona*, has been elucidated by using RNAseq technique (Horie T. et al., 2018). Single cell technology can thus be used to reconstruct the developmental patterns of the Central Nervous System. Moreover, the outputs of scRNAseq coupled with CRISPR/Cas9 technology can greatly improve our understanding about the interaction amongst the identified genes during specification and differentiation of the tissues of interest.

Although most of the processes regarding early neural development and some of the related gene regulatory pathways have been identified (Satoh et al., 2003; Sasakura et al., 2012), the results of the studies on scRNAseq of neural tissues of *Ciona* represent a great and encouraging starting point to continue using this technique to thoroughly depict the single cell transcriptome profiles, since early stages of *Ciona* embryogenesis in order to finely reconstruct, step by step, the developmental patterning of the Central Nervous System.

ATAC–seq

Animal development is driven by changes in the gene regulatory networks, which play a key role in the evolution of the animal body plans. The ATAC–seq represents a powerful technique for the identification and study of chromatin accessibility. Applying this technique in embryos, at different developmental stages and on different cell types, permits to investigate the regulatory pathways in the regions of interest (Magri et al., 2019). ATAC-seq has been recently developed for *Ciona* to profile chromatin accessibility through transitions from mesoderm to distinct fate-restricted heart and pharyngeal muscle precursors (Racioppi et al., 2019a).

These recent and extraordinary advantages in biological techniques, together with the affirmation of the tunicate as simple chordate model organism, are growingly increasing the possibilities to shed light on the multiple aspects and functions of the genes and their regulatory pathways.

CONCLUSION

In the course of evolution, increasingly intricate gene regulatory networks have been elaborated to orchestrate the specification, patterning and differentiation of diverse cell types in order to build up complex and different organisms. Ascidians, with their unique phylogenetic position, together with the simplicity of their typical chordate body plan and structures, represent for more than a century a powerful model organism to approach "simplified versions" of complex biological mechanisms. In these organisms it is possible to study gene regulatory networks, with a cellular resolution unprecedented in chordate models. Furthermore, the adaptation to this model system of the most updated technologies and high-throughput strategies for targeted loss-of-function and whole genome analyses, coupled with the use of live imaging approaches and computational methods, are more and more permitting to get important insights on developmental strategies exploited during chordate evolution. A clear example is just represented by the sensory organs, for which we are close to depict the developmental programs adopted by pigmented and photoreceptor cell lineages, at single cell level, starting from the early developmental stages, within the frame of whole CNS development. Under this perspective, the single cell transcriptomic analyses and the data already available will be instrumental to depict the set of transcripts of each lineage in order to better define the evolutionary relationships (i) of *Ciona* PR complexes with the photoreceptor organs of other chordates and (ii) of BV with vertebrate midbrain, with the aim to get further insights on the appearance and modifications of these structures in the course of evolution.

Collectively, these studies will permit to reveal the basic developmental mechanisms for forming chordate pigment and photoreceptor cells and the common/different evolutionary strategies adopted by higher vertebrates to build more complex structures.

Furthermore, the application of the most cutting-edge techniques can be extended to other biological fields besides developmental biology, as evolution and ecology, with the purpose to deepen the knowledge in the emergent discipline of Eco-evo-devo.

AUTHOR CONTRIBUTIONS

AP and PO: reviewing bibliography and drafting the manuscript. AS and FR: drafting and revising the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

PO has been supported by SZN OU Ph.D. fellowship. AP has been supported by SZN/University of Naples Federico II Veterinary Sciences Ph.D. fellowship. Funding for this review are from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730984, ASSEMBLE Plus project.

REFERENCES

Abitua, P. B., Wagner, E., Navarrete, I. A., and Levine, M. (2012). Identification of a rudimentary neural crest in a non-vertebrate chordate. *Nature* 492, 104–107. doi: 10.1038/nature11589

Arshavsky, V. Y., Lamb, T. D., and Pugh, E. N. Jr. (2002). G proteins and phototransduction. *Annu. Rev. Physiol.* 64, 153–187.

Bailey, T. J., El-Hodiri, H., Zhang, L., Shah, R., Mathers, E. H., and Jamrich, M. (2004). Regulation of vertebrate eye development by Rx genes. *Int. J. Dev. Biol.* 48, 761–770. doi: 10.1387/ijdb.041878tb

Bates, W. R. (1995). Direct development in the ascidian *Molgula retortiformis* (Verrill, 1871). *Biol. Bull.* 188, 16–22. doi: 10.2307/1542063

Bates, W. R., and Mallett, J. E. (1991). Ultrastructural and histochemical study of anural development in the ascidian *Molgula pacifica* (Huntsman). *Rouxs. Arch. Dev. Biol.* 200, 193–201. doi: 10.1007/bf00361337

Berrill, N. J., and Watson, D. M. S. (1931). VI. studies in tunicate development, part II.abbreviation of development in the molgulidæ. *Philos. Trans. R. Soc. London Series B Containing Papers Biol. Character* 219, 281–346. doi: 10.1098/rstb.1931.0006

Blomhoff, R., and Blomhoff, H. K. (2006). Overview of retinoid metabolism and function. *J. Neurobiol.* 66, 606–630. doi: 10.1002/neu.20242

Bostwick, M., Smith, E. L., Borba, C., Newman-Smith, E., Guleria, I., Kourakis, M. J., et al. (2020). Antagonistic inhibitory circuits integrate visual and gravitactic behaviors. *Curr. Biol.* 30, 600–609.e2.

Bouchemousse, S., Liautard-Haag, C., Bierne, N., and Viard, F. (2016). Distinguishing contemporary hybridization from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated Ciona species. *Mol. Ecol.* 25, 5527–5542. doi: 10.1111/mec.13854

Brunetti, R., Gissi, C., Pennati, R., Caicci, F., Gasparini, F., and Manni, L. (2015). Morphological evidence that the molecularly determined Ciona intestinalis type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. *J. Zoolog. Syst. Evol. Res.* 53, 186–193. doi: 10.1111/jzs.12101

Cao, C., Lemaire, L. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., et al. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571, 349–354. doi: 10.1038/s41586-019-1385-y

Caputi, L., Andreakis, N., Mastrototaro, F., Cirino, P., Vassillo, M., and Sordino, P. (2007). Cryptic speciation in a model invertebrate chordate. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9364–9369. doi: 10.1073/pnas.0610158104

Caracciolo, A., Gesualdo, I., Branno, M., Aniello, F., Di Lauro, R., and Palumbo, A. (1997). Specific cellular localization of tyrosinase mRNA during *Ciona intestinalis* larval development. *Dev. Growth Differ.* 39, 437–444. doi: 10.1046/j.1440-169x.1997.t01-3-00004.x

Cole, A. G., and Meinertzhagen, I. A. (2004). The central nervous system of the ascidian larva: mitotic history of cells forming the neural tube in late embryonic *Ciona intestinalis*. *Dev. Biol.* 271, 239–262. doi: 10.1016/j.ydbio.2004.04.001

Coppola, U., Kamal, A. K., Stolfi, A., and Ristoratore, F. (2020). The cis-regulatory code for kelch-like 21/30 specific expression in *Ciona robusta* sensory organs. *Front. Cell Dev. Biol.* 8:569601.

D'Aniello, E., Pezzotti, M. R., Locascio, A., and Branno, M. (2011). Onecut is a direct neural-specific transcriptional activator of Rx in *Ciona intestinalis*. *Dev. Biol.* 355, 358–371. doi: 10.1016/j.ydbio.2011.05.584

D'Aniello, S., D'Aniello, E., Locascio, A., Memoli, A., Corrado, M., Russo, M. T., et al. (2006). The ascidian homolog of the vertebrate homeobox gene Rx is essential for ocellus development and function. *Differentiation* 74, 222–234. doi: 10.1111/j.1432-0436.2006.00071.x

Darras, S., and Nishida, H. (2001). The BMP/CHORDIN antagonism controls sensory pigment cell specification and differentiation in the ascidian embryo. *Dev. Biol.* 236, 271–288. doi: 10.1006/dbio.2001.0339

Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298, 2157–2167. doi: 10.1126/science.1080049

Eakin, R. M., and Kuda, A. (1971). Ultrastructure of sensory receptors in *Ascidian tadpoles*. *Z. Zellforsch. Mikrosk. Anat.* 112, 287–312. doi: 10.1007/bf02584045

Esposito, R., D'Aniello, S., Squarzoni, P., Pezzotti, M. R., Ristoratore, F., and Spagnuolo, A. (2012). New insights into the evolution of metazoan tyrosinase gene family. *PLoS One* 7:e35731. doi: 10.1371/journal.pone.0035731

Gainous, T. B., Wagner, E., and Levine, M. (2015). Diverse ETS transcription factors mediate FGF signaling in the Ciona anterior neural plate. *Dev. Biol.* 399, 218–225. doi: 10.1016/j.ydbio.2014.12.032

Gandhi, S., Haeussler, M., Razy-Krajka, F., Christiaen, L., and Stolfi, A. (2017). Evaluation and rational design of guide RNAs for efficient CRISPR/Cas9-mediated mutagenesis in Ciona. *Dev. Biol.* 425, 8–20. doi: 10.1016/j.ydbio.2017.03.003

Gorman, A. L. F., McReynolds, J. S., and Barnes, S. N. (1971). Photoreceptors in primitive chordates: fine structure, hyperpolarizing receptor potentials, and evolution. *Science* 172, 1052–1054. doi: 10.1126/science.172.3987.1052

Haupaix, N., Abitua, P. B., Sirour, C., Yasuo, H., Levine, M., and Hudson, C. (2014). Ephrin-mediated restriction of ERK1/2 activity delimits the number of pigment cells in the Ciona CNS. *Dev. Biol.* 394, 170–180. doi: 10.1016/j.ydbio.2014.07.010

Horie, R., Hazbun, A., Chen, K., Cao, C., Levine, M., and Horie, T. (2018). Shared evolutionary origin of vertebrate neural crest and cranial placodes. *Nature* 560, 228–232. doi: 10.1038/s41586-018-0385-7

Horie, T., Horie, R., Chen, K., Cao, C., Nakagawa, M., Kusakabe, T. G., et al. (2018). Regulatory cocktail for dopaminergic neurons in a protovertebrate identified by whole-embryo single-cell transcriptomics. *Genes Dev.* 1, 1297–1302. doi: 10.1101/gad.317669.118

Horie, T., Nakagawa, M., Orii, H., and Tsuda, M. (2002). Whole structure of the photoreceptors in the ascidian larva visualized by an antibody against arrestin (Ci-Arr). *J. Photosci.* 9, 272–274.

Horie, T., Sakurai, D., Ohtsuki, H., Terakita, A., Shichida, Y., Usukura, J., et al. (2008). Pigmented and nonpigmented ocelli in the brain vesicle of the ascidian larva. *J. Comp. Neurol.* 509, 88–102. doi: 10.1002/cne.21733

Huber, J. L., da Silva, K. B., Bates, W. R., and Swalla, B. J. (2000). The evolution of anural larvae in molgulid ascidians. *Semin. Cell Dev. Biol.* 11, 419–426. doi: 10.1006/scdb.2000.0195

Hudson, C., Lotito, S., and Yasuo, H. (2007). Sequential and combinatorial inputs from nodal, Delta2/Notch and FGF/MEK/ERK signalling pathways establish a grid-like organisation of distinct cell identities in the ascidian neural plate. *Development* 134, 3527–3537. doi: 10.1242/dev.002352

Imai, K. S., Hino, K., Yagi, K., Satoh, N., and Satou, Y. (2004). Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development* 131, 4047–4058. doi: 10.1242/dev.01270

Imai, K. S., Stolfi, A., Levine, M., and Satou, Y. (2009). Gene regulatory networks underlying the compartmentalization of the *Ciona central* nervous system. *Development* 136, 285–293. doi: 10.1242/dev.026419

Jeffery, W. R., and Swalla, B. J. (1990). The myoplasm of ascidian eggs: a localized cytoskeletal domain with multiple roles in embryonic development. *Semin. Cell Biol.* 1, 373–381.

Jeffery, W. R., Swalla, B. J., Ewing, N., and Kusakabe, T. (1999). Evolution of the ascidian anural larva: evidence from embryos and molecules. *Mol. Biol. Evol.* 16, 646–654. doi: 10.1093/oxfordjournals.molbev.a026147

Jiang, D., Tresser, J. W., Horie, T., Tsuda, M., and Smith, W. C. (2005). Pigmentation in the sensory organs of the ascidian larva is essential for normal behavior. *J. Exp. Biol.* 208, 433–438. doi: 10.1242/jeb.01420

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Katz, M. J. (1983). Comparative anatomy of the tunicate tadpole. *Ciona intestinalis*. *Biol. Bull.* 164, 1–27. doi: 10.2307/1541186

Kim, K., Gibboney, S., Razy-Krajka, F., Lowe, E. K., Wang, W., and Stolfi, A. (2020). Regulation of neurogenesis by FGF signaling and neurogenin in the invertebrate chordate ciona. *Front. Cell Dev. Biol.* 8:477.

Kourakis, M. J., Borba, C., Zhang, A., Newman-Smith, E., Salas, P., Manjunath, B., et al. (2019). Parallel visual circuitry in a basal chordate. *Elife* 8:e44753. doi: 10.7554/eLife.44753

Kusakabe, T., Kusakabe, R., Kawakami, I., Satou, Y., Satoh, N., and Tsuda, M. (2001). Ci-opsin1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian *Ciona intestinalis*. *FEBS Lett.* 506, 69–72. doi: 10.1016/s0014-5793(01)02877-0

Kusakabe, T., Swalla, B. J., Satoh, N., and Jeffery, W. R. (1996). Mechanism of an evolutionary change in muscle cell differentiation in ascidians with different modes of development. *Dev. Biol.* 174, 379–392. doi: 10.1006/dbio.1996.0082

Lacaze-Duthiers, H., and de Lacaze-Duthiers, H. (1858). *Histoire de l'organisation, du Développement, des Moeurs et des Rapports Zoologiques du Dentale*. Paris: Librairie de Victor Masson Rue de l'ecole de medecine.

Levy, C., Khaled, M., and Fisher, D. E. (2006). MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol. Med.* 12, 406–414. doi: 10.1016/j.molmed.2006.07.008

Magri, M. S., Jiménez-Gancedo, S., Bertrand, S., Madgwick, A., Escrivà, H., Lemaire, P., et al. (2019). Assaying chromatin accessibility using ATAC-seq in invertebrate chordate embryos. *Front. Cell Dev. Biol.* 7:372.

Mast, S. O. (1921). Reactions to light in the larvae of the ascidians, *Amaroucium constellatum* and *Amaroucium pellucidum* with special reference to photic orientation. *J. Exp. Zool.* 34, 148–187. doi: 10.1002/jez.1400340204

Meinertzhagen, I. A., Lemaire, P., and Okamura, Y. (2004). The neurobiology of the ascidian tadpole larva: recent developments in an ancient chordate. *Annu. Rev. Neurosci.* 27, 453–485. doi: 10.1146/annurev.neuro.27.070203.144255

Miya, T., Morita, K., Suzuki, A., Ueno, N., and Satoh, N. (1997). Functional analysis of an ascidian homologue of vertebrate Bmp-2/Bmp-4 suggests its role in the inhibition of neural fate specification. *Development* 124, 5149–5159. doi: 10.1242/dev.124.24.5149

Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., et al. (1996). XTcf-3 transcription factor mediates beta-catenin-induced axis formation in *Xenopus embryos*. *Cell* 86, 391–399. doi: 10.1016/s0092-8674(00)80112-9

Moret, F., Christiaen, L., Deyts, C., Blin, M., Joly, J.-S., and Vernier, P. (2005). The dopamine-synthesizing cells in the swimming larva of the tunicateCiona intestinalisare located only in the hypothalamus-related domain of the sensory vesicle. *Eur. J. Neurosci.* 21, 3043–3055. doi: 10.1111/j.1460-9568.2005.04147.x

Nakashima, Y., Kusakabe, T., Kusakabe, R., Terakita, A., Shichida, Y., and Tsuda, M. (2003). Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and two putative visual cycle proteins are expressed in whole brain of a primitive chordate. *J. Comp. Neurol.* 460, 180–190. doi: 10.1002/cne.10645

Nicol, D., and Meinertzhagen, I. A. (1991). Cell counts and maps in the larval central nervous system of the ascidian *Ciona intestinalis* (L.). *J. Comp. Neurol.* 309, 415–429. doi: 10.1002/cne.903090402

Nishida, H. (1987). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. up to the tissue restricted stage. *Dev. Biol.* 121, 526–541. doi: 10.1016/0012-1606(87)90188-6

Nishida, H., and Satoh, N. (1989). Determination and regulation in the pigment cell lineage of the ascidian embryo. *Dev. Biol.* 132, 355–367. doi: 10.1016/0012-1606(89)90232-7

Nishiyama, A., and Fujiwara, S. (2008). RNA interference by expressing short hairpin RNA in the *Ciona intestinalis* embryo. *Dev. Growth Differ.* 50, 521–529. doi: 10.1111/j.1440-169x.2008.01039.x

Okamura, Y., Nishino, A., Murata, Y., Nakajo, K., Iwasaki, H., Ohtsuka, Y., et al. (2005). Comprehensive analysis of the ascidian genome reveals novel insights into the molecular evolution of ion channel genes. *Physiol. Genomics* 22, 269–282. doi: 10.1152/physiolgenomics.00229.2004

Oonuma, K., and Kusakabe, T. G. (2019). Spatio-temporal regulation of Rx and mitotic patterns shape the eye-cup of the photoreceptor cells in Ciona. *Dev. Biol.* 445, 245–255. doi: 10.1016/j.ydbio.2018.11.011

Oonuma, K., Tanaka, M., Nishitsuji, K., Kato, Y., Shimai, K., and Kusakabe, T. G. (2016). Revised lineage of larval photoreceptor cells in *Ciona reveals* archetypal collaboration between neural tube and neural crest in sensory organ formation. *Dev. Biol.* 420, 178–185. doi: 10.1016/j.ydbio.2016.10.014

Pan, Y., Martinez-De Luna, R. I., Lou, C.-H., Nekkalapudi, S., Kelly, L. E., Sater, A. K., et al. (2010). Regulation of photoreceptor gene expression by the retinal homeobox (Rx) gene product. *Dev. Biol.* 339, 494–506. doi: 10.1016/j.ydbio.2009.12.032

Racioppi, C., Coppola, U., Christiaen, L., and Ristoratore, F. (2019b). Transcriptional regulation of Rab32/38, a specific marker of pigment cell formation in *Ciona robusta*. *Dev. biol.* 448, 111–118. doi: 10.1016/j.ydbio.2018.11.013

Racioppi, C., Kamal, A. K., Razy-Krajka, F., Gambardella, G., Zanetti, L., di Bernardo, D., et al. (2014). Fibroblast growth factor signalling controls nervous system patterning and pigment cell formation in *Ciona intestinalis*. *Nat. Commun.* 5:4830.

Racioppi, C., Valoroso, M. C., Coppola, U., Lowe, E. K., Titus Brown, C., Swalla, B. J., et al. (2017). Evolutionary loss of melanogenesis in the tunicate Molgula occulta. *EvoDevo* 8:11. doi: 10.1186/s13227-017-0074-x

Racioppi, C., Wiechecki, K. A., and Christiaen, L. (2019a). Combinatorial chromatin dynamics foster accurate cardiopharyngeal fate choices. *eLife* 8:e49921. doi: 10.7554/elife.49921

Razy-Krajka, F., Brown, E. R., Horie, T., Callebert, J., Sasakura, Y., Joly, J.-S., et al. (2012). Monoaminergic modulation of photoreception in ascidian: evidence for a proto-hypothalamo-retinal territory. *BMC Biol.* 10:45. doi: 10.1186/1741-7007-10-45

Roux, C., Tsagkogeorga, G., Bierne, N., and Galtier, N. (2013). Crossing the species barrier: genomic hotspots of introgression between two highly divergent *Ciona intestinalis* species. *Mol. Biol. Evol.* 30, 1574–1587. doi: 10.1093/molbev/mst066

Ryan, K., Lu, Z., and Meinertzhagen, I. A. (2016). The CNS connectome of a tadpole larva of *Ciona intestinalis* (L.) highlights sidedness in the brain of a chordate sibling. *Elife* 5:e16962. doi: 10.7554/eLife.16962

Sakurai, D., Goda, M., Kohmura, Y., Horie, T., Iwamoto, H., Ohtsuki, H., et al. (2004). The role of pigment cells in the brain of ascidian larva. *J. Comput. Neurol.* 12, 70–82. doi: 10.1002/cne.20142

Salas, P., Vinaithirthan, V., Newman-Smith, E., Kourakis, M. J., and Smith, W. C. (2018). Photoreceptor specialization and the visuomotor repertoire of the primitive chordate Ciona. *J. Exp. Biol.* 221:jeb177972. doi: 10.1242/jeb.177972

Sasakura, Y., Mita, K., Ogura, Y., and Horie, T. (2012). Ascidians as excellent chordate models for studying the development of the nervous system during embryogenesis and metamorphosis. *Dev. Growth Differ.* 54, 420–437. doi: 10.1111/j.1440-169x.2012.01343.x

Sato, A., Satoh, N., and Bishop, J. D. D. (2012). Field identification of "types" A and B of the ascidian *Ciona intestinalis* in a region of sympatry. *Mar. Biol.* 159, 1611–1619. doi: 10.1007/s00227-012-1898-5

Sato, S., and Yamamoto, H. (2001). Development of pigment cells in the brain of ascidian tadpole larvae: insights into the origins of vertebrate pigment cells. *Pigment Cell Res.* 14, 428–436. doi: 10.1034/j.1600-0749.2001.140602.x

Satoh, N., Satou, Y., Davidson, B., and Levine, M. (2003). Ciona intestinalis: an emerging model for whole-genome analyses. *Trends Genet.* 19, 376–381. doi: 10.1016/s0168-9525(03)00144-6

Sharma, S., Wang, W., and Stolfi, A. (2019). Single-cell transcriptome profiling of the *Ciona larval* brain. *Dev. Biol.* 448, 226–236. doi: 10.1016/j.ydbio.2018.09.023

Shimeld, S. M., and Levin, M. (2006). Evidence for the regulation of left-right asymmetry in *Ciona intestinalis* by ion flux. *Dev. Dyn.* 235, 1543–1553. doi: 10.1002/dvdy.20792

Squarzoni, P., Parveen, F., Zanetti, L., and Ristoratore, F. (2011). FGF/MAPK/Ets signaling renders pigment cell precursors competent to respond to Wnt signal by directly controlling Ci-Tcf transcription. *Development* 138, 1421–1432. doi: 10.1242/dev.057323

Stolfi, A., Gandhi, S., Salek, F., and Christiaen, L. (2014). Tissue-specific genome editing in *Ciona embryos* by CRISPR/Cas9. *Development* 141, 4115–4120. doi: 10.1242/dev.114488

Svane, I., and Young, C. M. (1989). The ecology and behaviour of ascidian larvae. *Oceanogr. Mar. Biol.* 27, 45–90.

Swalla, B. J., and Jeffery, W. R. (1990). *Modification of the Myoplasm in Eggs of Anural Ascidians. in American Zoologist (Amer Soc Zoologists 1041 New Hampshire St, Lawrence, Ks 66044)*. American Society Zoologists.

Swalla, B. J., and Jeffery, W. R. (1992). Modification of the myoplasm in eggs of anural ascidians. *American Zoologist* 30, A42–A42.

Tagawa, K., Jeffery, W. R., and Satoh, N. (1997). The recently-described ascidian species *Molgula tectiformis* is a direct developer. *Zoolog. Sci.* 14, 297–303. doi: 10.2108/zsj.14.297

Tassy, O., Daian, F., Hudson, C., Bertrand, V., and Lemaire, P. (2006). A quantitative approach to the study of cell shapes and interactions during early chordate embryogenesis. *Curr. Biol.* 16, 345–358. doi: 10.1016/j.cub.2005.12.044

Tief, K., Schmidt, A., Aguzzi, A., and Beermann, F. (1996). Tyrosinase is a new marker for cell populations in the mouse neural tube. *Dev. Dyn.* 205, 445–456. doi: 10.1002/(sici)1097-0177(199604)205:4<445::aid-aja8>3.0.co;2-i

Tsuda, M., Sakurai, D., and Goda, M. (2003). Direct evidence for the role of pigment cells in the brain of ascidian larvae by laser ablation. *J. Exp. Biol.* 206(Pt. 8), 1409–1417. doi: 10.1242/jeb.00235

Wagner, E., and Levine, M. (2012). FGF signaling establishes the anterior border of the *Ciona neural* tube. *Development* 139, 2351–2359. doi: 10.1242/dev.078485

Whittaker, J. R. (1979). Development of vestigial tail muscle acetylcholinesterase in embryos of an anural ascidian species. *Biol. Bull.* 156, 393–407. doi: 10.2307/1540926

Yoshida, R., Sakurai, D., Horie, T., Kawakami, I., Tsuda, M., and Kusakabe, T. (2004). Identification of neuron-specific promoters in *Ciona intestinalis*. *Genesis* 39, 130–140. doi: 10.1002/gene.20032

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Olivo, Palladino, Ristoratore and Spagnuolo. 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) and the copyright owner(s) 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 Cell and Developmental Biology

REVIEW
published: 28 September 2021
doi: 10.3389/fcell.2021.760366

Developmental Evolution of Hypaxial Muscles: Insights From Cyclostomes and Chondrichthyans

Rie Kusakabe[1]*, Masako Tanaka[1] and Shigeru Kuratani[1,2]

[1] Laboratory for Evolutionary Morphology, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Japan,
[2] Evolutionary Morphology Laboratory, RIKEN Cluster for Pioneering Research (CPR), Kobe, Japan

OPEN ACCESS

Edited by:
Salvatore D'Aniello,
Zoological Station Anton Dohrn, Italy

Reviewed by:
Renata Freitas,
Universidade do Porto, Portugal
Sebastian Shimeld,
University of Oxford, United Kingdom

***Correspondence:**
Rie Kusakabe
rie.kusakabe@riken.jp

Specialty section:
This article was submitted to
Evolutionary Developmental Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 18 August 2021
Accepted: 10 September 2021
Published: 28 September 2021

Citation:
Kusakabe R, Tanaka M and
Kuratani S (2021) Developmental
Evolution of Hypaxial Muscles:
Insights From Cyclostomes
and Chondrichthyans.
Front. Cell Dev. Biol. 9:760366.
doi: 10.3389/fcell.2021.760366

Jawed vertebrates possess two distinct groups of muscles in the trunk (epaxial and hypaxial muscles) primarily defined by the pattern of motor innervation from the spinal cord. Of these, the hypaxial group includes muscles with highly differentiated morphology and function, such as the muscles associated with paired limbs, shoulder girdles and tongue/infrahyoid (hypobranchial) muscles. Here we summarize the latest findings on the evolutionary mechanisms underlying the morphological variety of hypaxial musculature, with special reference to the molecular insights obtained from several living species that diverged early in vertebrate evolution. Lampreys, extant jawless vertebrates, lack many of derived traits characteristic of the gnathostomes, such as jaws, paired fins and epaxial/hypaxial distinction of the trunk skeletal musculatures. However, these animals possess the primitive form of the hypobranchial muscle. Of the gnathostomes, the elasmobranchs exhibit developmental mode of hypaxial muscles that is not identical to that of other gnathostomes in that the muscle primordia relocate as coherent cell aggregates. Comparison of expression of developmental genes, including *Lbx* genes, has delineated the temporal order of differentiation of various skeletal muscles, such as the hypobranchial, posterior pharyngeal and cucullaris (trapezius) muscles. We have proposed that the sequential addition of distal muscles, associated with expression of duplicated *Lbx* genes, promoted the elaboration of skeletal musculature. These analyses have revealed the framework of an evolutionary pathway that gave rise to the morphological complexity and diversity of vertebrate body patterns.

Keywords: vertebrates, lamprey, shark, skeletal muscle, hypobranchial muscles, fin muscles

INTRODUCTION

Among different types of vertebrate muscles, skeletal muscles are those connected to skeletal elements to exert the force required for all kinds of movement. To achieve locomotion, respiration, nutrition uptake and even communication among individuals, development of the skeletal muscles must be precisely controlled to be positioned along the axes of the body. Histologically, the vertebrate muscles are also categorized into two major groups; striated (skeletal and cardiac) and smooth muscles. Although the contractile apparatus of all these muscle types utilizes actomyosin, which originated in early eukaryotes, bilateral paraxial muscle derivatives, all of which are skeletal muscles, are considered a chordate innovation, as deduced from the absence of homologous muscle in ambulacrarians (echinoderms and hemichordates; Brunet and Arendt, 2016; Brunet et al., 2016; Inoue and Satoh, 2018). In the vertebrate trunk, all the skeletal muscles are derived from somites, the segmented units of bilateral paraxial mesoderm aligned along the midline. Somites contain precursor of skeletal muscles, axial skeletons and connective tissues (tendons and ligaments) that differentiate in a coordinated manner.

Amniotes such as mammals, birds and reptiles exhibit a complex combination of skeletal musculature in the trunk (**Figure 1A**). Muscle layers at the dorsal and the ventral sides of the body are categorized into epaxial and hypaxial muscles, each innervated by dorsal and ventral rami of spinal motor nerves, respectively (**Figure 1A**). Epaxial and hypaxial distinction is more conspicuous in cartilaginous and bony fish; the two compartments are partitioned by the horizontal myoseptum, a connective tissue sheet located at either side of the notochord (**Figure 1B**). In the tail region, epaxial and hypaxial portions occupy mirror-imaged bulks across the horizontal myoseptum. Their coordinated contraction is suitable for S-shaped swimming movement (Nair et al., 2015 and references therein).

In contrast to those in fish, the epaxial muscles in amniotes comprise of minor members of trunk skeletal muscles, i.e., the intrinsic back muscles connecting the vertebral columns. On the other hand, the hypaxial portion includes a variety of body wall muscles, limb muscles, some members of shoulder girdle muscles such as the trapezius (cucullaris) muscles and tongue and infrahyoid muscle (**Figures 1A**, **2A,C**). Tongue and infrahyoid muscles are collectively called hypobranchial muscles (HBMs), as they develop from a bilateral pair of muscle primordia extending rostrally from the anterior somites (for details, see below). Hypaxial somites also contribute to mammalian diaphragm, a component of the respiratory system. Thus, epaxial/hypaxial muscle distribution in amniotes appear suitable for the terrestrial life; epaxial muscles are specialized for maintenance of posture, whereas hypaxial muscles are responsible for locomotion in the terrestrial environment, prey-capturing, and respiration.

The epaxial/hypaxial distinction of skeletal muscles would have been acquired early in vertebrate evolution, but it is not observed in the cyclostome species, the lampreys and hagfish. Lamprey trunk muscle consists of concentric layers of stacks of muscle sheets that are not compartmented in dorsoventral direction (**Figure 1C**). The muscle fibers are aligned parallel to the body axis with both ends attached to the myosepta of chevron-shaped myotomes (Peters and Mackay, 1961). This configuration appears similar to the somatic musculature of amphioxus, a non-vertebrate chordate (**Figure 1D**), which is composed of flat lamellae (Peachey, 1961). Thus, epaxial/hypaxial distinction of skeletal muscles would have been acquired in the common ancestor of jawed vertebrates (gnathostomes) after the divergence of the cyclostomes (Kuraku and Kuratani, 2006; **Figure 1**). In this review, we summarize the current understanding of the developmental pathways and genetic control in action during myogenesis in lampreys and sharks and discuss about the developmental mechanisms that contributed to the elaborate structure of the skeletal muscles found in extant vertebrates.

DEVELOPMENT OF THE TRUNK SKELETAL MUSCLES

The vertebrate skeletal muscles develop exclusively from mesodermal tissues of the embryo. In the trunk, all the skeletal muscles originate from dermomyotomes, the myogenic compartment of the somites (**Figure 2A**). During early development, skeletal muscles originate from the dorsal and ventral lips of the dermomyotomes, each of which is the source of epaxial and hypaxial muscles, respectively. The dermomyotomal lips serve as “growth zones” for the expanding trunk musculature, as they add actively proliferating myogenic cells to myotomes that develop medially (Hollway and Currie, 2003). Some hypaxial muscles undergo delamination from the ventral lips and are directed distally to travel a long distance to the sites of differentiation, where they fuse to form multinuclear myofibers (**Figure 2A**). These cells remain undifferentiated and mesenchymal during the migration. This type of muscle differentiation occurs in the limbs, tongue and mammalian diaphragm, all of which are organs distally located from the axis or even in the head. These distally relocating hypaxial muscle precursor cells have been called “migratory muscle precursors” (MMPs) as opposed to the other myotomal cells also formed from dermomyotomes but non-migratory (Alvares et al., 2003).

Molecular mechanisms regulating the migration of MMPs have been best studied in paired limb development in mouse and chick (Christ and Ordahl, 1995; Birchmeier and Brohmann, 2000; **Figure 2A**). *Pax3*, a gene encoding a paired class of homeobox transcription factor, is initially expressed in the whole dermomyotome, but secondarily becomes restricted to hypaxial muscle progenitors. At limb and occipital levels of the trunk, *Pax3*-positive cells also express *Lbx1*, another gene encoding a homeodomain transcription factor, required for delamination and migration of MMPs. MMPs also express receptor proteins-encoding genes such as *c-met* and *CXCR4*, whose products are detected and bound by specific ligands, HGF and Sdf1, respectively, distributed in the mesodermal mesenchyme filling the limb bud (Dietrich, 1999; Dietrich et al., 1999; Vasyutina et al., 2005). Interactions between these receptors and ligands lead MMPs to the correct location to form muscles.

In anamniotes, such as amphibians and fish, skeletal muscles of the trunk differentiate *in situ* from the bulk of myogenic cells initially occupying a large portion of the somites. Nevertheless, a thin lateral layer of epithelium that covers the myotome has been discovered in these animals (Kusakabe and Kuratani, 2005; Scaal and Wiegreffe, 2006). In teleosts, this thin layer expresses *Pax3* gene and contribute myogenic cells to the pectoral and pelvic fins. *Lbx1* gene is expressed in the progenitors

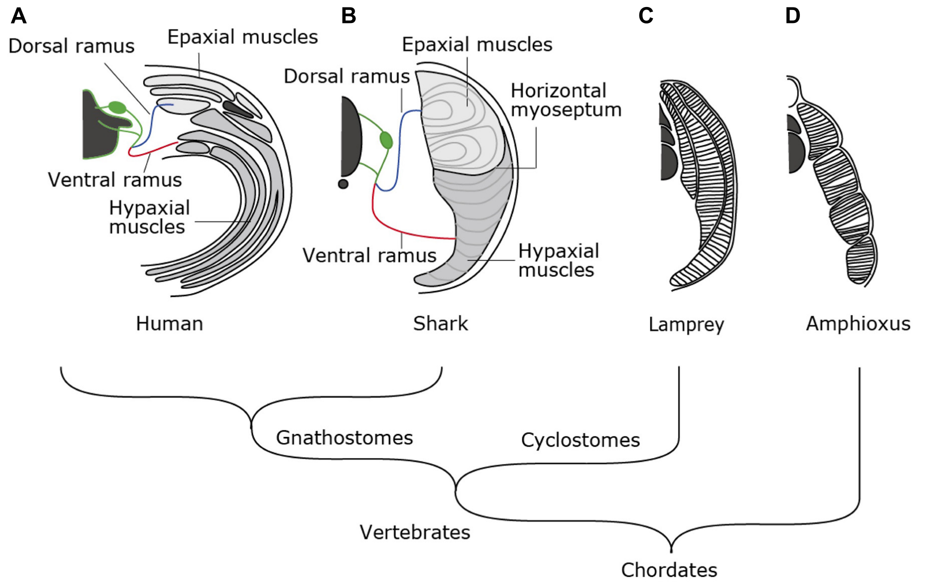


FIGURE 1 | Morphological patterns of trunk skeletal muscles. In amniotes **(A)**, trunk skeletal muscles consist of multiple muscle layers categorized into epaxial or hypaxial muscle, innervated by dorsal or ventral ramus of spinal nerves, respectively. **(B)** Epaxial and hypaxial domains are clearly segregated by horizontal myoseptum in cartilaginous and bony fishes. **(C)** Lampreys have concentric layers of muscle tissues that are not compartmentalized into epaxial/hypaxial groups. **(D)** Amphioxus exhibits lamella-like body wall muscles consisting of layers of myofiber sheets.

of paired fin muscles, as well as the sternohyoid and the posterior hypaxial muscles, both of which cover the ventrolateral aspects of the anterior larval trunk (Sagarin et al., 2019). Likewise, amphibian hypaxial muscles, including those in the limbs that develop at metamorphosis, develop from the dermomyotome-like layer covering the embryonic myotome (Martin and Harland, 2006).

SKELETAL MUSCULATURE OF ADULT LAMPREYS AND SHARK

As mentioned above, body wall muscles of cyclostomes are not segregated into epaxial/hypaxial domains. Cyclostomes also lack the development of paired fins similarly to amphioxus (**Figures 1C,D**, **2B**). However, lampreys possess bilateral HBMs that cover ventrolateral wall of the pharynx (reviewed in Kusakabe and Kuratani, 2020; **Figure 2B**). Lamprey HBMs are innervated by hypoglossal nerve which is homologous to the 12th cranial nerve of mammals. This innervation pattern, as well as the developmental process described below, supports the homology of lamprey HBMs to the tongue and infrahyoid muscles which derive from the hypaxial portion of anterior somites (Kuratani, 1997; Kusakabe and Kuratani, 2005; for the hagfish, see Oisi et al., 2015). Lampreys also possess the supra- and infraoptic muscles derived from the most anterior somites and they extend into the rostral head (**Figure 2B**). The infraoptic muscle is innervated by a part of the spinal nerve plexus located caudal to the vagal (X) motorneuron, suggesting its possible linkage with trapezius muscle of the amniote neck (Tada and Kuratani, 2015). Collectively, although lampreys lack epaxial/hypaxial distinction of trunk skeletal muscles, their myotomes exhibit characteristics similar to the specialized hypaxial muscles found in amniotes.

Chondrichthyans, such as the shark, have served as models for development of early vertebrates (Boisvert et al., 2019; **Figure 2C**). In particular, chondrichthyan HBMs have been intensively studied with anatomical and developmental viewpoints (Edgeworth, 1935; Miyake et al., 1992; Kusakabe et al., 2020). In sharks and rays, the hypobranchial group of muscles has been documented as somite derivatives with variable names depending on the species (Edgeworth, 1935). In the catshark, HBMs are segregated into the posterior bilateral domain (coracoarcualis muscle, CAC; also called rectus cervicus) and the medially located anterior domain (coracomandibularis, CMD; also called geniohyoideus; Miyake et al., 1992; Kusakabe et al., 2020). Cucullais muscles are also conspicuous in these animals; it appears as a thin single muscle plate covering the dorsal aspect of the posterior branchial region. In skate, cucullaris muscles are innervated by the vagus as well as the rostral spinal nerves (Tanaka, 1988; Boord and Sperry, 1991).

Chondrichthyans have also attracted attention concerning their paired fins musculature (Goodrich, 1930; Ziermann et al., 2017; Turner et al., 2019). Skates and rays have evolved extraordinarily broad paired fins that generate strong forward propulsion, whereas sharks exhibit narrower pectoral and pelvic fins separated by a broad flank region. Based on a series of classical observations performed on catshark *Schyliorhinus canicula,* shark pectoral fin muscles have been believed to emerge as a direct extension of myotome cells into fin bud, a developmental mode suggested to be primitive (Galis, 2001; Hollway and Currie, 2003). This view has also been supported

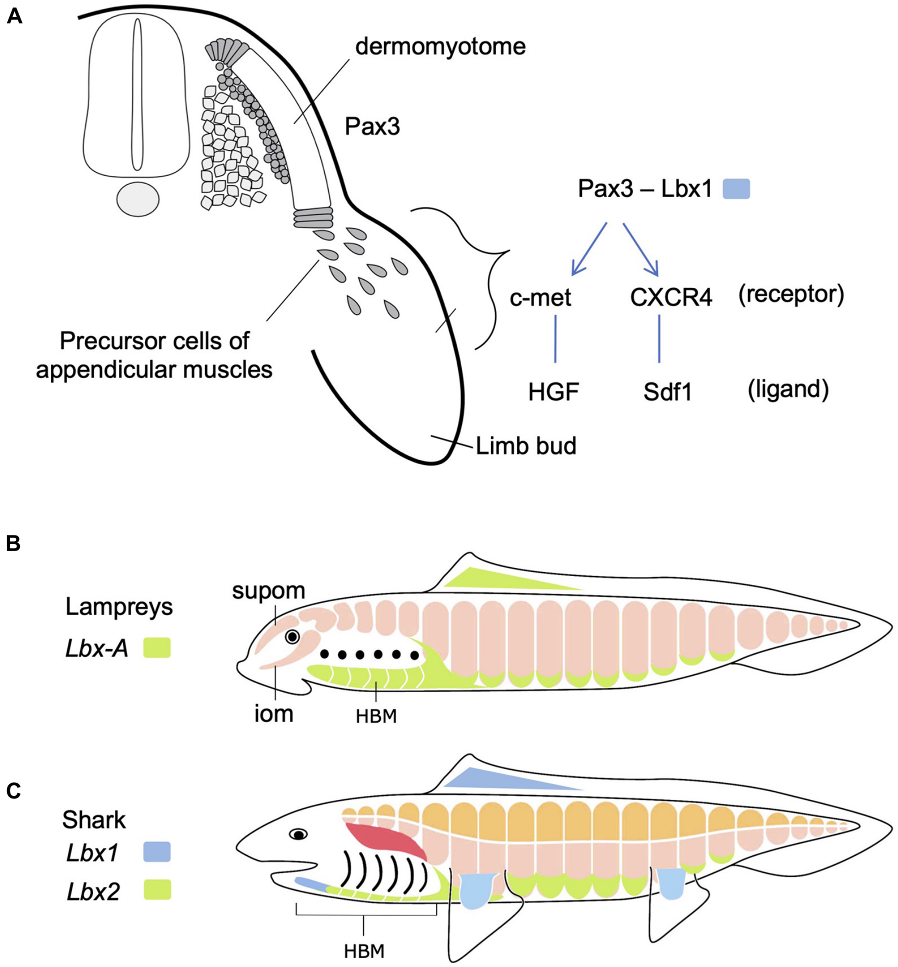


FIGURE 2 | Development of hypaxial muscles controlled by specific regulatory genes. **(A)** In amniotes, *Pax3*-positive dermomyotome supplies precursor cells of epaxial and hypaxial muscles. At the limb levels, precursor cells of appendicular muscles delaminate and migrate distally in the limb bud. In these cells, *Lbx*1 is specifically expressed. These cells also express c-met and CXCR4 receptors that detect ligands in the limb bud mesenchyme. See text for details. **(B)** Updated and modified scheme from Kusakabe and Kuratani (2005), based on Kusakabe et al. (2020). Lampreys have no epaxial/hypaxial distinction, no paired appendages, but possess hypobranchial muscle (HBM) that require the unique *Lbx* gene during development. iom, infraoptic muscle; supom, supraoptic muscle. **(C)** On the other hand, gnathostomes (e.g., shark) have epaxial (orange) vs. hypaxial (pink) distinction of skeletal muscles and paired appendages. *Lbx* paralogs *Lbx1* (blue) and *Lbx2* (green) are differentially expressed in many of the hypaxial muscles. Developmental origin of shoulder girdle muscles, such as cucullaris (red), remains controversial (see also **Figure 5**).

by nerve supply to the paired fins; each myotome extending into the fin bud receives motor neurons from the ventral roots of the neighboring spinal nerves (Goodrich, 1930). However, recent progress of molecular analysis in chondrichthyan embryos provided important clues that their paired fin muscles would develop with a similar developmental mechanism to that in MMPs in amniote limbs, as detailed below (Okamoto et al., 2019; Turner et al., 2019; Kusakabe et al., 2020).

DEVELOPMENT OF LAMPREY HYPOBRANCHIAL MUSCLE

During lamprey embryogenesis, although the earliest differentiation of somitic skeletal muscles takes place at early neurula stage, the HBMs differentiates as late as pre-ammocoete larval stage (stage 29; Tahara, 1988; Kusakabe et al., 2020). HBMs first appear as bilateral rods lying on the ventrolateral aspects of the pharynx, spanning between the velum and the pericardium, being segmented in correspondence to the pharyngeal arches, not to the myotomes (Kuratani et al., 1999; **Figure 3**). Prior to differentiation of HBMs, at stage 28, the hypoglossal nerve (XII) extends its axons, circumventing the posterior edge of the pharyngeal region (Tada and Kuratani, 2015; **Figure 3A**).

Extension of hypoglossal axons matches well with the distribution of precursor cells of HBMs that leave anterior somites, proceed initially ventrally, and then anteriorly along the ventral floor of the pharynx. The streams of these myogenic cells are comparable with the hypoglossal cord, the tightly condensed

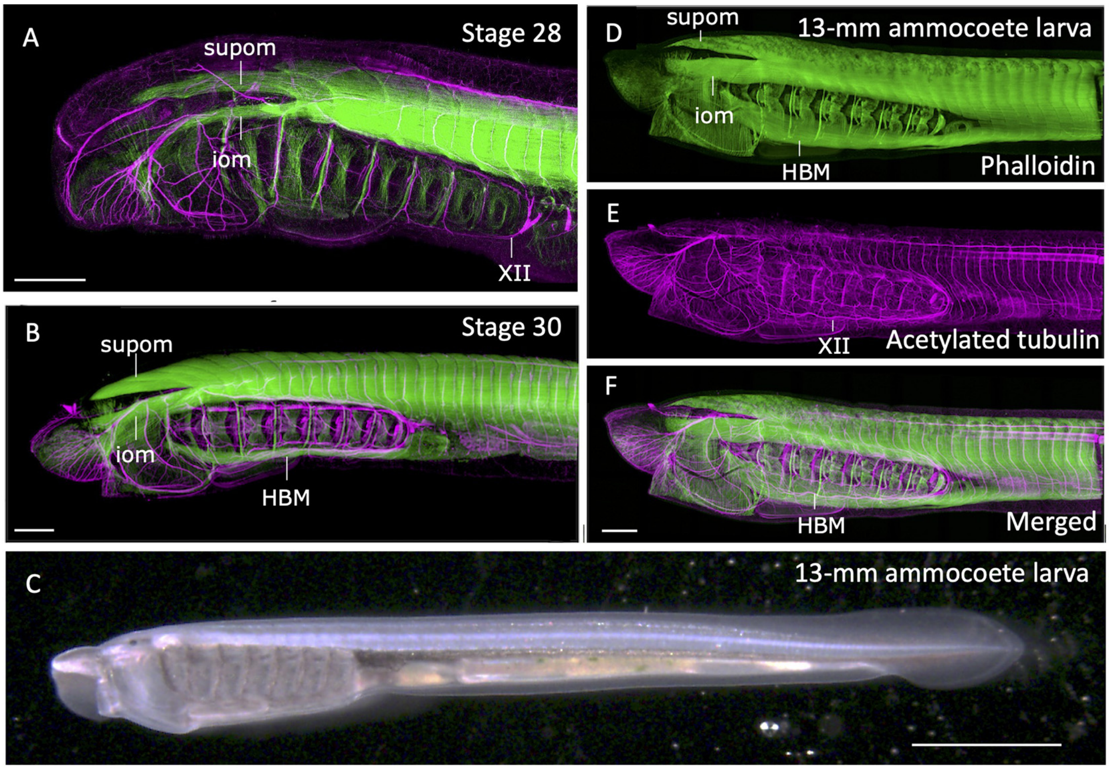


FIGURE 3 | Formation of somitic muscles during lamprey development. The embryos and larva were stained with phalloidin (green; skeletal muscles) and acetylated tubulin antibody (magenta; neurons) as described in Kusakabe et al. (2020). **(A)** Head region of a stage 28. Extending hypoglossal nerve (XII) can be observed, but no hypobranchial muscle (HBM) has yet to be differentiated. **(B)** Stage 30 pre-ammocoete embryo. Initial differentiation of HBM is visible. **(C)** A 13-mm ammocoete larva in bright field viewed from left side. **(D–F)** Anterior region of the larva in **(C)** stained with phalloidin and acetylated tubulin antibody. HBM has undergone growth in width, and the XII nerve has fully extended its axon into the oral region. iom, infraoptic muscle; supom, supraoptic muscle. Scale bars: 2 mm **(C)**, 0.2 mm **(A,B,F)**.

strand of myoblasts giving rise to tongue and infrahyoid muscles of amniotes (Kuratani et al., 1999; for hagfish, see Oisi et al., 2015; reviewed in Kusakabe and Kuratani, 2020). These cells express *Pax37A* and *LbxA* genes, the cognate genes of amniote *Pax3* and *Lbx1*, respectively (Kusakabe et al., 2011). These genes are initially expressed broadly in somitic mesoderm, but become intense in the hypoglossal cord as it extends anteriorly. Targeted deletion of *LbxA* gene leads to the specific deficiency of HBMs, suggesting the molecular mechanism involving *Lbx* gene is required for the formation of HBMs, as is the case in amniote MMP-derived muscles (Brohmann et al., 2000; Kusakabe et al., 2020).

At the dorsal region of the head, infra- and supraopitic muscles are readily observed as phalloidin-positive anterior myotomes directly extending in chevron-shape as early as stage 26 (Kuratani et al., 1999; Kusakabe et al., 2020; **Figure 3A**). In the growing ammocoete larvae, these muscles, together with late-differentiating HBM, expand in width and cover the entire head (**Figures 3B–F**). Unlike the tongue muscle of the amniotes, HBM of the lamprey do not fuse in the midline but remain in the lateral pharyngeal wall. Caudally the HBMs are continuous to the body wall muscle, which form as the ventral edges of the trunk somites extending toward ventral midline (Kusakabe et al., 2020; **Figures 3D,F**). Thus, the lamprey somitic muscles as a whole can be viewed as the simplified "hypaxial" muscles of vertebrates in the absence of paired appendages, but associated with MMP-like developmental pattern.

DEVELOPMENT OF THE SHARK SKELETAL MUSCLES

In early development of a catshark *Schyliorhinus torazame*, skeletal muscle differentiation is first observed in trunk somites aligned bilaterally along the axis (**Figures 4A–C**). Unlike that of the lamprey, early shark somites exhibit conspicuous epaxial/hypaxial domains segregated by horizontal myoseptum (arrowheads in **Figure 4**). As the pectoral and pelvic fin buds become prominent (stage 28 and onward; Ballard et al., 1993), the flank region exhibits secondary muscle differentiation as a ventral extension of the neighboring hypaxial myotomes (**Figure 4E**). Each segment of this muscle extension has a gradient in maturation of fibers, as shown by immunoreactivity to myosin heavy chain antibody, in a way that muscle differentiation proceeds from anterior-to-posterior direction (**Figure 4E**). Distribution pattern of the differentiated myofibers is complementary to the expression pattern of *Lbx2* gene, the paralog of *Lbx1* of the shark, at slightly earlier stage (detailed below; Kusakabe et al., 2020; **Figure 4D**). These

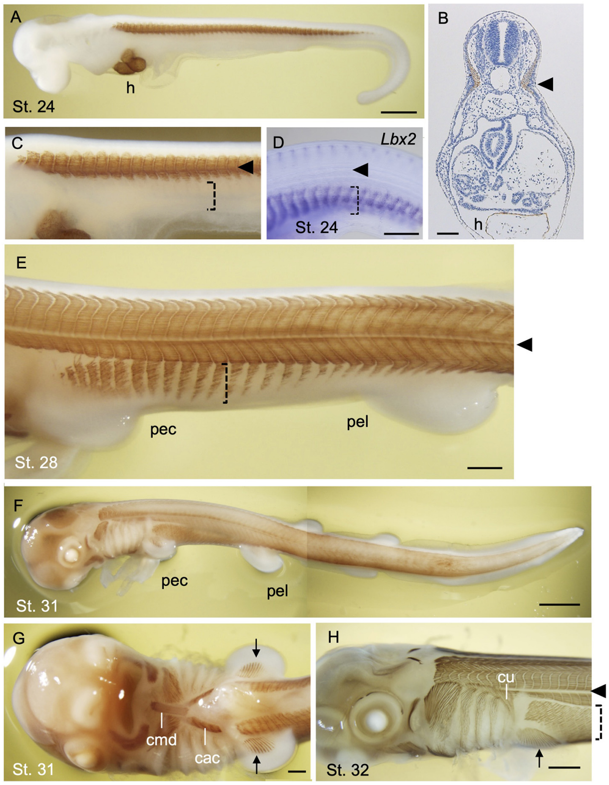


FIGURE 4 | Development of skeletal muscles during catshark development. The embryos were stained with the myosin heavy chain antibody (Brown; **A–C,E–H**) and antisense probe for *Lbx2* mRNA **(D)** as described in Kusakabe et al. (2020). Black arrowhead indicates the boundary of epaxial and hypaxial domains. **(A)** A stage 24 embryo. Muscle differentiation starts at trunk somites. The heart (h) is also stained. **(B)** A transverse section of stage 24 embryo at the level of the heart stained with myosin heavy chain antibody. **(C)** Magnified view of **(A)**. Body wall muscle has yet to be differentiated (dotted bracket). **(D)** *Lbx2* gene expression is observed in undifferentiated body wall muscle primordia (dotted bracket). **(E)** A stage 28 embryo. Body wall muscle is now undergoing differentiation (dotted bracket). **(F)** A stage 31 embryo. **(G)** Ventral view of the embryo in **(F)**. Coracomandibularis (cmd) and coracoarcualis (cac) muscles consist the hypobranchial muscles of catshark. Pectoral fin muscles are also visible (arrow). **(H)** A stage 32 embryo associate viewed from the left. cac, coracoarcualis muscle; cmd, coracomandibularis muscle; cu, cucullaris muscle; h, heart; pec, pectoral fin; pel, pelvic fin. Scale bars: 2 mm **(F)**, 1 mm **(A,H)**, 0.5 mm **(D,E,G)**, 0.1 mm **(B)**.

later differentiating body wall myofibers will form the rectus abdominus muscle that extends to meet at the ventral midline (**Figures 4F,H**).

Differentiated muscles in the pectoral fin buds are observed as late as stage 31 (**Figures 4F,G**), followed by those observed in pelvic fin buds at stage 32 (data not shown) Delayed differentiation of pelvic fin muscles is also the case in another shark species *Schyliorhinus canicula* (Ziermann et al., 2017), as well as in many of the teleosts in which pelvic fin development occurs postembryonically (Parichy et al., 2009). As mentioned above, shark fin muscles had long been thought to emerge as a direct extension of neighboring epithelial myotomes and thus called "muscle bud" by Goodrich (1930). Recently, however, molecular evidence from sharks and skates showed that "muscle bud" in the paired fins express *Pax3* and *Lbx1* genes, and that muscle precursors detach from the dermomyotomes at the base of the fin bud and move distally before the onset of muscle differentiation (Okamoto et al., 2017; Turner et al., 2019; Kusakabe et al., 2020). This implies that shark fin muscles are regulated by a genetic mechanism common to the amniote MMPs described above. It is noteworthy, however, that muscle precursors of shark paired fins leave the ventral lip of dermomyotomes as cell aggregates, not as dispersed mesenchymal cells, unlike MMPs in amniote limb bud (Kusakabe et al., 2020). Accordingly, these aggregates of myogenic cells are positive for ZO-1 antibody, an indicator of epithelial characteristics (Mayeuf-Louchart et al., 2016; Kusakabe et al., 2020).

In addition to the appendicular muscles, shark HBMs also exhibit more characteristics in common with the amniotes, rather than to that of the lamprey. Initial differentiation of HBM can be observed as a pair of muscle fibers flanking the heart (Kusakabe et al., 2020). These rods of skeletal muscle are a derivative of anterior somites and express *Lbx2* (Kusakabe et al., 2020) and give rise to the CAC muscle (**Figure 4G**). The most anterior tips of forming CAC join at the ventral midline and connect to the CMD muscle, which appears later than CAC (Kusakabe et al., 2020; **Figure 4G**). CMD muscle originates from the precursor cells at the tip of the bilateral CAC which is marked by expression of *Lbx1* gene (Kusakabe et al., 2020). CAC and CMD muscles of the shark, together with laterally oriented CHY muscles (see Kusakabe et al., 2020), show conspicuous homology to those of amphibians as well as of the amniote tongue and infrahyoid muscles, with respect to the anatomical structure and developmental gene expression (Miyake et al., 1992).

Another member of the hypaxial muscles found in the late shark embryo is the cucullaris muscle (**Figure 4H**). It differentiates as a triangular muscle located at the dorsal side of the posterior pharyngeal arches, inserting into the scapulocoracoid catilage. It has been shown in skates that cucullaris muscles are innervated by both vagus (X) and rostral spinal nerves (Boord and Sperry, 1991), which is a characteristic of the neck muscles of amniotes. Thus, developmental patterns of HBM and fin-associated muscles collectively suggest the ancestral status of distal hypaxial formation in chondrichthyans with respect to the vertebrate evolution.

DISCUSSION

The epaxial/hypaxial distinction of skeletal muscles would have appeared in the ancestral vertebrate, primitively associated with a simple locomotive action driven by the somitic musculature. Spatial segregation and mutually independent innervation of the two domains might have facilitated the adaptation to the terrestrial habitat of tetrapods—epaxial muscles have become specialized for the protection of the internal organs and maintenance of posture, whereas hypaxial muscles become highly functionally specialized to support a wide variety of locomotive, respiratory and feeding movement. Consistently, during the

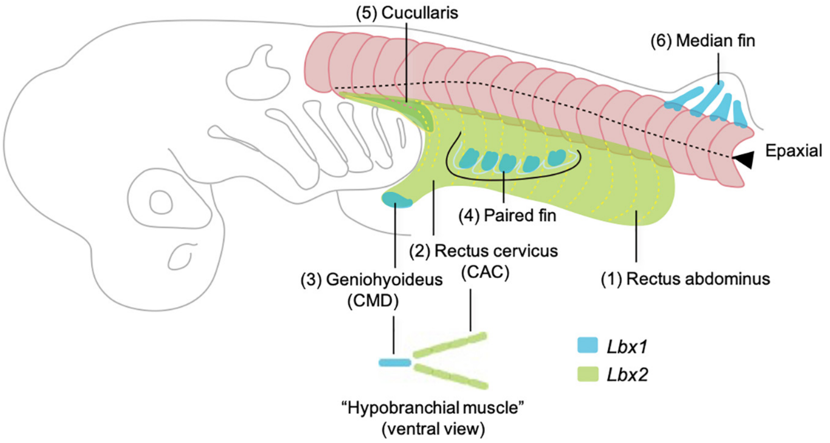


FIGURE 5 | Developmental order of somitic skeletal muscles with emphasis on the embryonic origins based on the observation in Kusakabe et al. (2020). Here the muscles are named in generalized terms for vertebrate anatomy based on Miyake et al. (1992). Numbers (1–6) indicate the temporal order of muscle differentiation. Body wall muscle (1, rectus abdominus) and the posterior portion of hypobranchial muscles (2, rectus cervicus, or coracoarcualis) express *Lbx2* and are possibly homologous to the lamprey body wall muscle and hypobranchial muscles, respectively. *Lbx1* regulates the distal hypobranchial muscle (geniohyoideus, or coracomandibularis) and both lateral and median fins.

fin-to-limb transition, a relative quantity of locomotor muscles, all of which belong to hypaxial muscles, drastically increased (Mansuit and Herrel, 2021).

The timing of acquisition of HBM, which develops as a part of hypaxial muscles in gnathostomes, remains unclear—it could precede the appearance of epaxial/hypaxial boundary, as HBM is present in the cyclostomes (reviewed in Kusakabe and Kuratani, 2020). With reference to the behavior of muscle precursor cells, Burke and Nowicki (2003) proposed a different context for categorization of somite derivatives. According to their scheme, the somites are divided into primaxial and abaxial domains. The abaxial group consists of muscles primarily categorized into the hypaxial and also associated with MMP-like mode of development (Nowicki et al., 2003); that is, the ventral lips of dermomyotomes at specific anteroposterior levels undergo delamination and long-distance migration, and become associated with skeletons and tendons of non-somitic origins. The primaxial muscles, on the other hand, consist of all epaxial muscles and non-migratory population of hypaxial muscle, both of which become associated with somitic skeletons.

Primaxial/abaxial distinction well explains the developmental mode of complex hypaxial musculature in gnathostomes. In abaxial muscles, such as appendicular muscles, shoulder girdle muscles and the ventral abdominal muscles, prolonged expression of *Pax3* and *Lbx* genes are correlated with the maintenance of proliferating muscle progenitors during the relocation from the paraxial somites. In this regard, HBMs also can be categorized as abaxial—they undergo long-distance extension from the ventral portion of anterior somites, and differentiate into mature muscle fibers in non-somitic, environment (cephalic neural crest cells and pharyngeal mesoderm; reviewed in Kusakabe and Kuratani, 2020). In the trunk, the boundary of primaxial/abaxial domains, termed lateral somitc frontier (LSF), exists at the dorsal edge of the lateral plate mesoderm. LSF-equivalent can also be defined for HBM development which involves extension of hypoglossal cord into the pharyngeal region. This boundary would overlap with S-shaped head-trunk interface, on which a variety of neck/shoulder muscles differentiate under the influence from neural crest cells giving rise to cephalic cartilages and connective tissues (Kuratani, 1997).

Shark fin muscles have long been thought to develop as a direct extension of epithelial myotomes and thus to represent the primitive mode of appendicular muscle formation (Goodrich, 1930; Neyt et al., 2000; Galis, 2001; Hollway and Currie, 2003), a concept contradicting with the acquisition of *Pax3*- and *Lbx*-positive HBM in the common ancestor of cyclostomes and gnathostomes (Kusakabe et al., 2011). Recent insights from catshark have provided a clearer view of the evolutionary events that occurred during the establishment of hypaxial muscles in early vertebrates (**Figure 5**). *Lbx2*-positive hypaxial muscles (equivalent to rectus abdominus and rectus cervicus) would represent the ancestral abaxial muscle which is equivalent to the lamprey body wall muscle and HBM. *Lbx1*-positive muscles in the shark (equivalent to geniohyoideus and fin muscles) would represent the anatomical modification which occurred in the distal somitic muscles that underlying the complex skeletal musculature of gnathostomes. These observations have led to a novel evolutionary hypothesis in which the ancestor of vertebrates acquired a primitive version of hypobranchial muscle that represents the advanced mode of somitic muscle development (Kusakabe et al., 2020).

AUTHOR CONTRIBUTIONS

RK and SK discussed the data and wrote the manuscript. MT contributed to production of the figure panels and proofread the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

RK was supported by the Grant-in-Aid for Scientific Research (C) (19K06683) and Takeda Science Foundation. SK was supported by the Grant-in-Aid for Scientific Research on Innovative Areas (Research in a Proposed Research Area) (17H06385).

ACKNOWLEDGMENTS

We express our gratitude to Yayoi Nakai for her technical assistance and to all other members of Laboratory for Evolutionary Morphology at RIKEN BDR for valuable discussions.

REFERENCES

Alvares, L. E., Schubert, F. R., Thorpe, C., Mootoosamy, R. C., Cheng, L., Parkyn, G., et al. (2003). Intrinsic, Hox-dependent cues determine the fate of skeletal muscle precursors. *Dev. Cell* 5, 379–390. doi: 10.1016/s1534-5807(03)00263-6

Ballard, W. W., Mellinger, J., and Lechenault, H. (1993). A Series of normal stages for development of *Scyliorhinus-canicula*, the lesser spotted dogfish (*Chondrichthyes*, *Scyliorhinidae*). *J. Exp. Zool.* 267, 318–336. doi: 10.1002/jez.1402670309

Birchmeier, C., and Brohmann, H. (2000). Genes that control the development of migrating muscle precursor cells. *Curr. Opin. Cell Biol.* 12, 725–730. doi: 10.1016/S0955-0674(00)00159-9

Boisvert, C. A., Johnston, P., Trinajstic, K., and Johanson, Z. (2019). "Chondrichthyan evolution, diversity, and senses," in *Heads, Jaws, and Muscles: Anatomical, Functional, and Developmental Diversity in Chordate Evolution*, eds J. M. Ziermann, R.E. Diaz Jr, and R. Diogo (Cham: Springer International Publishing), 65–91.

Boord, R. L., and Sperry, D. G. (1991). Topography and nerve supply of the cucullaris (trapezius) of skates. *J. Morphol.* 207, 165–172. doi: 10.1002/jmor.1052070207

Brohmann, H., Jagla, K., and Birchmeier, C. (2000). The role of Lbx1 in migration of muscle precursor cells. *Development* 127, 437–445.

Brunet, T., and Arendt, D. (2016). From damage response to action potentials: early evolution of neural and contractile modules in stem eukaryotes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150043. doi: 10.1098/rstb.2015.0043

Brunet, T., Fischer, A. H., Steinmetz, P. R., Lauri, A., Bertucci, P., and Arendt, D. (2016). The evolutionary origin of bilaterian smooth and striated myocytes. *Elife* 5:e19607. doi: 10.7554/eLife.19607

Burke, A. C., and Nowicki, J. L. (2003). A new view of patterning domains in the vertebrate mesoderm. *Dev. Cell* 4, 159–165. doi: 10.1016/S1534-5807(03)00033-9

Christ, B., and Ordahl, C. P. (1995). Early stages of chick somite development. *Anat. Embryol. (Berl)* 191, 381–396.

Dietrich, S. (1999). Regulation of hypaxial muscle development. *Cell Tissue Res.* 296, 175–182.

Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E., Yamaai, T., et al. (1999). The role of SF/HGF and c-Met in the development of skeletal muscle. *Development* 126, 1621–1629.

Edgeworth, F. H. (1935). *The Cranial Muscles of Vertebrates*. London: Cambridge University Press.

Galis, F. (2001). Evolutionary history of vertebrate appendicular muscle. *Bioessays* 23, 383–387. doi: 10.1002/bies.1056

Goodrich, E. S. (1930). *Studies on the Structure and Development of Vertebrates*. London: Macmillan and Co., Ltd.

Hollway, G. E., and Currie, P. D. (2003). Myotome meanderings. Cellular morphogenesis and the making of muscle. *EMBO Rep.* 4, 855–860. doi: 10.1038/sj.embor.embor920

Inoue, J., and Satoh, N. (2018). Deuterostome genomics: lineage-specific protein expansions that enabled chordate muscle evolution. *Mol. Biol. Evol.* 35, 914–924. doi: 10.1093/molbev/msy002

Kuraku, S., and Kuratani, S. (2006). Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zoolog. Sci.* 23, 1053–1064. doi: 10.2108/zsj.23.1053

Kuratani, S. (1997). Spatial distribution of postotic crest cells defines the head/trunk interface of the vertebrate body: embryological interpretation of peripheral nerve morphology and evolution of the vertebrate head. *Anat. Embryol. (Berl)* 195, 1–13.

Kuratani, S., Horigome, N., and Hirano, S. (1999). Developmental morphology of the head mesoderm and reevaluation of segmental theories of the vertebrate head: evidence from embryos of an agnathan vertebrate, Lampetra japonica. *Dev. Biol.* 210, 381–400. doi: 10.1006/dbio.1999.9266

Kusakabe, R., Higuchi, S., Tanaka, M., Kadota, M., Nishimura, O., and Kuratani, S. (2020). Novel developmental bases for the evolution of hypobranchial muscles in vertebrates. *BMC Biol.* 18:120. doi: 10.1186/s12915-020-00851-y

Kusakabe, R., Kuraku, S., and Kuratani, S. (2011). Expression and interaction of muscle-related genes in the lamprey imply the evolutionary scenario for vertebrate skeletal muscle, in association with the acquisition of the neck and fins. *Dev. Biol.* 350, 217–227. doi: 10.1016/j.ydbio.2010.10.029

Kusakabe, R., and Kuratani, S. (2005). Evolution and developmental patterning of the vertebrate skeletal muscles: perspectives from the lamprey. *Dev. Dyn.* 234, 824–834. doi: 10.1002/dvdy.20587

Kusakabe, R., and Kuratani, S. (2020). "Development and evolution of the neck muscles," in *Evolutionary Developmental Biology: A Reference Guide*, eds L. Nuno de la Rosa and G. Müller (Cham: Springer International Publishing), 1–14.

Mansuit, R., and Herrel, A. (2021). The evolution of appendicular muscles during the fin-to-limb transition: possible insights through studies of soft tissues, a perspective. *Front. Ecol. Evol.* 9:508. doi: 10.3389/fevo.2021.702576

Martin, B. L., and Harland, R. M. (2006). A novel role for lbx1 in Xenopus hypaxial myogenesis. *Development* 133, 195–208. doi: 10.1242/dev.02183

Mayeuf-Louchart, A., Montarras, D., Bodin, C., Kume, T., Vincent, S. D., and Buckingham, M. (2016). Endothelial cell specification in the somite is compromised in Pax3-positive progenitors ofFoxc1/2conditional mutants, with loss of forelimb myogenesis. *Development* 143, 872–879. doi: 10.1242/dev.128017

Miyake, T., McEachran, J. D., and Hall, B. K. (1992). Edgeworth's legacy of cranial muscle development with an analysis of muscles in the ventral gill arch region of batoid fishes (Chondrichthyes: Batoidea). *J. Morphol.* 212, 213–256. doi: 10.1002/jmor.1052120304

Nair, A., Azatian, G., and McHenry, M. J. (2015). The kinematics of directional control in the fast start of zebrafish larvae. *J. Exp. Biol.* 218, 3996–4004. doi: 10.1242/jeb.126292

Neyt, C., Jagla, K., Thisse, C., Thisse, B., Haines, L., and Currie, P. D. (2000). Evolutionary origins of vertebrate appendicular muscle. *Nature* 408, 82–86. doi: 10.1038/35040549

Nowicki, J. L., Takimoto, R., and Burke, A. C. (2003). The lateral somitic frontier: dorso-ventral aspects of anterio-posterior regionalization in avian embryos. *Mech. Dev.* 120, 227–240. doi: 10.1016/s0925-4773(02)00415-x

Oisi, Y., Fujimoto, S., Ota, K. G., and Kuratani, S. (2015). On the peculiar morphology and development of the hypoglossal, glossopharyngeal and vagus nerves and hypobranchial muscles in the hagfish. *Zoolog. Lett.* 1:6. doi: 10.1186/s40851-014-0005-9

Okamoto, E., Kusakabe, R., Kuraku, S., Hyodo, S., Robert-Moreno, A., Onimaru, K., et al. (2017). Migratory appendicular muscles precursor cells in the common ancestor to all vertebrates. *Nat. Ecol. Evol.* 1, 1731–1736. doi: 10.1038/s41559-017-0330-4

Okamoto, E., Moriyama, Y., Kuraku, S., Kai, K. I., and Tanaka, M. (2019). Involvement of HGF/MET signaling in appendicular muscle development in cartilaginous fish. *Dev. Growth Differ.* 61, 97–103. doi: 10.1111/dgd.12591

Parichy, D. M., Elizondo, M. R., Mills, M. G., Gordon, T. N., and Engeszer, R. E. (2009). Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev. Dyn.* 238, 2975–3015. doi: 10.1002/dvdy.22113

Peachey, L. D. (1961). Structure of the longitudinal body muscles of amphioxus. *J. Biophys. Biochem. Cytol.* 10(Suppl), 159–176. doi: 10.1083/jcb.10.4.159

Peters, A., and Mackay, B. (1961). The structure and innervation of the myotomes of the lamprey. *J. Anat.* 95, 575–585.

Sagarin, K. A., Redgrave, A. C., Mosimann, C., Burke, A. C., and Devoto, S. H. (2019). Anterior trunk muscle shows mix of axial and appendicular developmental patterns. *Dev. Dyn.* 248, 961–968. doi: 10.1002/dvdy.95

Scaal, M., and Wiegreffe, C. (2006). Somite compartments in anamniotes. *Anat. Embryol. (Berl)* 211(Suppl. 1), 9–19. doi: 10.1007/s00429-006-0127-8

Tada, M. N., and Kuratani, S. (2015). Evolutionary and developmental understanding of the spinal accessory nerve. *Zoolog. Lett.* 1:4. doi: 10.1186/s40851-014-0006-8

Tahara, Y. (1988). Normal stages of development in the lamprey, *Lampetra reissneri* (Dybowski). *Zool. Sci.* 5, 109–118.

Tanaka, S. (1988). A macroscopical study of the trapezius muscle of sharks, with reference to the topographically related nerves and vein. *Anat. Anz.* 165, 7–21.

Turner, N., Mikalauskaite, D., Barone, K., Flaherty, K., Senevirathne, G., Adachi, N., et al. (2019). The evolutionary origins and diversity of the neuromuscular system of paired appendages in batoids. *Proc. Biol. Sci.* 286:20191571. doi: 10.1098/rspb.2019.1571

Vasyutina, E., Stebler, J., Brand-Saberi, B., Schulz, S., Raz, E., and Birchmeier, C. (2005). CXCR4 and Gab1 cooperate to control the development of migrating muscle progenitor cells. *Genes Dev.* 19, 2187–2198. doi: 10.1101/gad.346205

Ziermann, J. M., Freitas, R., and Diogo, R. (2017). Muscle development in the shark *Scyliorhinus canicula*: implications for the evolution of the gnathostome head and paired appendage musculature. *Front. Zool.* 14:31. doi: 10.1186/s12983-017-0216-y

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kusakabe, Tanaka and Kuratani. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 22 October 2021
doi: 10.3389/fcell.2021.744982

When Bigger Is Better: 3D RNA Profiling of the Developing Head in the Catshark *Scyliorhinus canicula*

Hélène Mayeur[1], Maxence Lanoizelet[1], Aurélie Quillien[2], Arnaud Menuet[3], Léo Michel[1], Kyle John Martin[4], Sébastien Dejean[5], Patrick Blader[2], Sylvie Mazan[1]*† and Ronan Lagadec[1]†

[1] CNRS, Sorbonne Université, UMR 7232-Biologie Intégrative des Organismes Marins (BIOM), Observatoire Océanologique, Banyuls sur Mer, France, [2] Molecular, Cellular and Developmental Biology (MCD UMR 5077), Centre de Biologie Intégrative (CBI, FR 3743), Université de Toulouse, CNRS, UPS, Toulouse, France, [3] UMR 7355, Experimental and Molecular Immunology and Neurogenetics, CNRS and University of Orléans, Orléans, France, [4] United Kingdom Research and Innovation, Biotechnology and Biological Sciences Research Council, Swindon, United Kingdom, [5] Institut de Mathématiques de Toulouse, Université de Toulouse, CNRS, UPS, UMR 5219, Toulouse, France

OPEN ACCESS

Edited by:
Salvatore D'Aniello,
Zoological Station Anton Dohrn, Italy

Reviewed by:
Linda Z. Holland,
University of California, San Diego,
United States
Daniel Medeiros,
University of Colorado, Boulder,
United States

***Correspondence:**
Sylvie Mazan
mazan@obs-banyuls.fr

†*These authors have contributed equally to this work and share last authorship*

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *21 July 2021*
Accepted: *28 September 2021*
Published: *22 October 2021*

Citation:
Mayeur H, Lanoizelet M, Quillien A, Menuet A, Michel L, Martin KJ, Dejean S, Blader P, Mazan S and Lagadec R (2021) When Bigger Is Better: 3D RNA Profiling of the Developing Head in the Catshark Scyliorhinus canicula. Front. Cell Dev. Biol. 9:744982. doi: 10.3389/fcell.2021.744982

We report the adaptation of RNA tomography, a technique allowing spatially resolved, genome-wide expression profiling, to a species occupying a key phylogenetic position in gnathostomes, the catshark *Scyliorhinus canicula*. We focused analysis on head explants at an embryonic stage, shortly following neural tube closure and of interest for a number of developmental processes, including early brain patterning, placode specification or the establishment of epithalamic asymmetry. As described in the zebrafish, we have sequenced RNAs extracted from serial sections along transverse, horizontal and sagittal planes, mapped the data onto a gene reference taking advantage of the high continuity genome recently released in the catshark, and projected read counts onto a digital model of the head obtained by confocal microscopy. This results in the generation of a genome-wide 3D atlas, containing expression data for most protein-coding genes in a digital model of the embryonic head. The digital profiles obtained for candidate forebrain regional markers along antero-posterior, dorso-ventral and left-right axes reproduce those obtained by *in situ* hybridization (ISH), with expected relative organizations. We also use spatial autocorrelation and correlation as measures to analyze these data and show that they provide adequate statistical tools to extract novel expression information from the model. These data and tools allow exhaustive searches of genes exhibiting any predefined expression characteristic, such a restriction to a territory of interest, thus providing a reference for comparative analyses across gnathostomes. This methodology appears best suited to species endowed with large embryo or organ sizes and opens novel perspectives to a wide range of evo-devo model organisms, traditionally counter-selected on size criterion.

Keywords: RNA tomography, catshark, forebrain patterning, correlation, auto-correlation

INTRODUCTION

Unraveling spatially resolved gene expression information is crucial for addressing a variety of developmental biology problems, including patterning, regional identity specification, and determination. These questions are currently revolutionized by emerging RNA-seq approaches and concomitantly developing analytical methods aimed at extracting relevant information from the resulting datasets (Lee, 2017; Sun et al., 2020; Waylen et al., 2020). One such approach is RNA tomography or tomo-seq, which was initially applied to the developing zebrafish and mouse (Junker et al., 2014). Briefly, this method relies on the sequencing of frozen sections along three orthogonal planes and subsequent reconstruction of a spatially resolved genome-wide RNA profile, using the iterative proportional fitting algorithm. This results in the generation of a digital 3D model of the tissue of interest, containing genome-wide expression information for every voxel of the model.

While the sequencing of sections has been successfully used for the identification of genes exhibiting regionalized expression along a polarity axis, 3D reconstructions using tomo-seq have only been obtained in the zebrafish and mouse thus far to our knowledge (Junker et al., 2014). We have tested the possibility to conduct this approach in a representative chondrichthyan, the catshark *Scyliorhinus canicula*. As a member of the sister group of osteichthyans, this species occupies a key phylogenetic position to address the origin of gnathostomes from an evo-devo perspective. Its analysis has thus provided insights into gnathostome ancestral features and gene regulatory network modifications, occurring concomitantly with the rise of gnathostome innovations, such as paired limbs, true teeth or articulated jaws, and helped to clarify ancestral features, when discrepancies are observed between actinopterygians and sarcopterygians (Coolen et al., 2008). It is also amenable to experimental approaches during development and while genome size (4.2 Gb) has long been an obstacle to whole genome sequencing projects, an annotated, chromosome-level genome *de novo* assembly has been recently released in this species.

We have focused on the embryonic head at stage 17, which shortly follows neural tube closure. A number of developmental processes of interest, such as the partitioning of the forebrain into its broad subdivisions (Santos-Durán et al., 2015), the appearance of the earliest diencephalic asymmetries (Lagadec et al., 2015), the specification of cephalic mesoderm components (Derobert et al., 2002), or the initiation of optic cup and otic placode formation (Plouhinec et al., 2005; O'Neill et al., 2007), indeed take place in the developing head at this stage. Here, we report the generation of a genome-wide 3D profile of the cephalic region at this stage, as well as the use of correlation and other statistical analysis tools to extract novel information from this model. In addition to providing reference expression data at the stage analyzed, this analysis validates the catshark as a relevant organism for tomo-seq characterizations of embryogenesis.

MATERIALS AND METHODS

Explant Dissection, Sectioning and RNA Extraction From Sections

Eggs from the catshark *S. canicula* were obtained from the Aquariology Service of the Banyuls sur Mer Oceanological Observatory. Stage 17 embryos (Ballard et al., 1993) were manually dissected and truncated anterior to the first pair of pharyngeal pouches. The cephalic explant was transferred to O.C.T. medium (Tissue-Tek O.C.T. compound, Sakura Finetek), frozen in liquid nitrogen prior to sectioning (18 mm cryostat sections). Each frozen section was transferred into Trizol (200 μL) immediately after collection. Total RNA was extracted using standard Trizol extraction protocol and purified on Macherey Nagel NucleoSpin RNA XS columns according to the manufacturer's instructions. 2 ml (1:200,000) spike-in RNA (ERCC) were added to each section prior RNA extraction as internal control.

Illumina Library Construction and Sequencing

cDNA synthesis, amplification, and Illumina library construction were conducted following (Holler and Junker, 2019) with the following modifications. Superscript II reverse-transcriptase was replaced by SuperScript IV (Invitrogen) and cDNA clean-up was done using Omega Bio-Tek's Mag-Bind TotalPure NGS. Sequencing was conducted on the DNB-seq platform (100 bp paired-end).

Read Mapping and Counting

Reads were mapped onto the reference database of predicted genes (provided in **Supplementary Table 1**; annotation in **Supplementary Table 2**) complemented with spike-in sequences. Briefly, the construction of this database (to be reported elsewhere) involved building an isoform collapsed version of the NCBI gene predictions annotated from the catshark genome assembly[1] and a subsequent extension of 3′ UTRs. The resulting gene models are referred to as genes hereafter for simplicity. Reference indexing, read mappings and quantifications were done using Kallisto v0.44.0 using the –bias parameter and bootstrapping 100 times. Read counts were aggregated for each gene and each sectioning plane, and normalized against the total spike-in read counts for each section. Ratios between spike-in input and output were calculated for the sum of all sections along each section plane.

3D Model Construction

To generate a 3D binary mask, a stage 17 head explant dissected as described above was fixed in PFA 4%, DAPI stained, dehydrated in methanol and cleared in Benzyl Alcohol/Benzyl Benzoate (1/2) prior to mounting and imaging with a SP8X confocal microscope (Leica). A 3D binary mask was built using Fiji to obtain a

---

[1] https://www.ncbi.nlm.nih.gov/assembly/GCF_902713615.1/#/st_Primary-Assembly

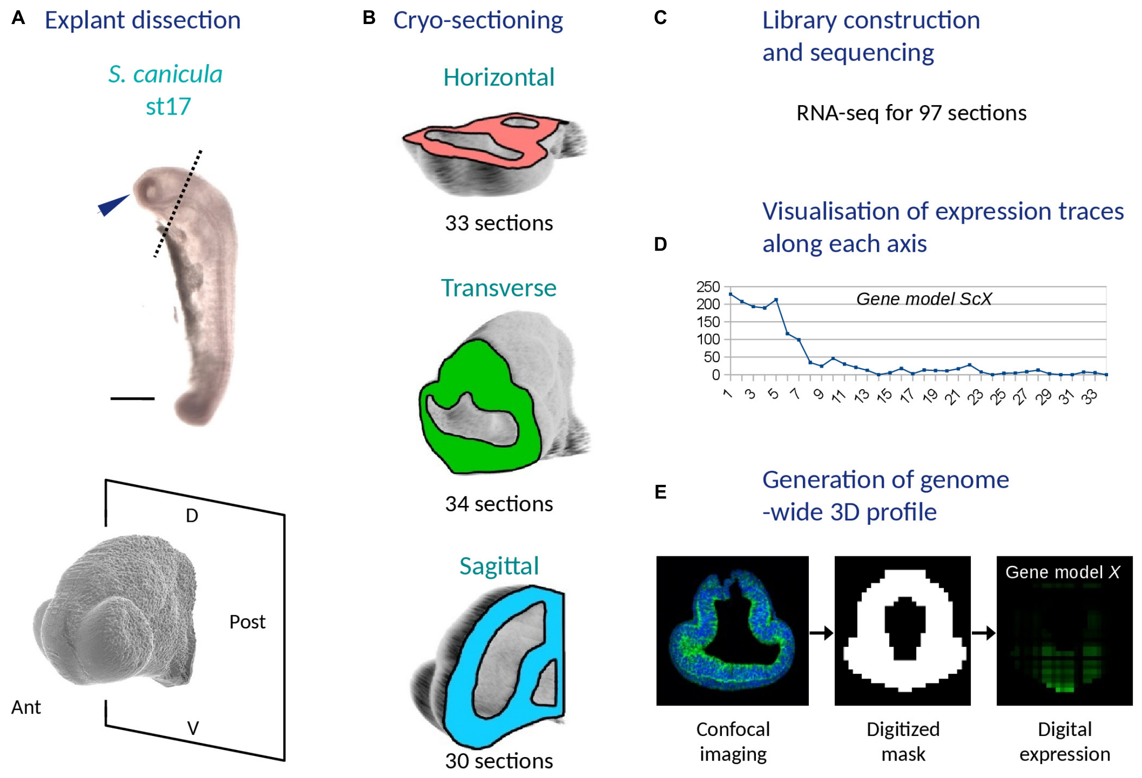

A
Explant dissection
S. canicula
st17
D
Post
Ant
V
B
Cryo-sectioning
Horizontal
33 sections
Transverse
34 sections
Sagittal
30 sections
C
Library construction
and sequencing
RNA-seq for 97 sections
D
Visualisation of expression traces
along each axis
Gene model ScX
250
200
150
100
50
0
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
E
Generation of genome
-wide 3D profile
Gene model X
Confocal
imaging
Digitized
mask
Digital
expression

FIGURE 1 | Main experimental steps used to generate a 3D RNA profile of the catshark embryonic head (stage 17). **(A)** Explant dissection: the top panel shows a left side photograph of a stage 17 catshark embryo with a blue arrowhead showing the dissected explant and a dotted line delimiting its posterior boundary and the bottom panel shows a higher magnification confocal view of the explant with the posterior dissection plane. **(B)** Cryo-sectioning: schemes show the planes used to generate series of cryo-sections (horizontal, pink; transverse, green; sagittal, blue). The total number of sections along each plane is indicated. **(C)** Library construction and sequencing resulting in RNA-seq data for a total of 97 sections. **(D)** Visualization of expression traces along each axis for all gene models. **(E)** Generation of a genome-wide 3D profile with its three steps (acquisition of a confocal 3D image of the explant, generation of a digitized model and computation of digital expressions for each voxel and gene model). D, dorsal; V, ventral; Ant, anterior; Post, posterior.

binary image of the tissue. It was then oriented transversally using Interactive Stack Rotation plugin and resized to match sectioning planes and section numbers along all three planes. The 3D expression genome-wide profile was reconstructed from normalized read counts for each gene and each section plane using the MATLAB code reported in Junker et al. (2014). In short, following the virtual partitioning of the binary mask in serial digital sections as described above, an iterative process was applied to 1D profiles of each gene in each sectioning plane, multiplying the 3D expression of the gene by said profile in each plane in succession. 3D profiles of selected genes were then visualized using Fiji. A custom Fiji macro was also written in order to get all 3D profiles, each viewed in all three sectioning planes, in the AVI format and with filenames indicating gene identification and annotation (.avi files for all genes available upon request).

Statistical Analysis

For the autocorrelation analysis, Moran's indexes were calculated for each gene on a volume defined as the voxels on adjacent sections on all three section planes, giving a 3 × 3 × 3 cube centered on each voxel examined. Supports for the indexes thus obtained were estimated by statistical tests, consisting in calculating the Pearson correlation between the expression of a gene between all possible pairs of neighboring voxels, using the R libraries spdep (especially moran.test) and ade4. Pearson correlations between the 3D expressions of each possible pair of genes in each voxel (18 μm cube at the intersection of one section of each sectioning plane) were calculated using the R command cor with default parameters.

In situ Hybridization of Whole-Mount Embryos and Sections

Whole-mount ISH and ISH of paraffin sections were conducted using standard protocols as described previously (Lagadec et al., 2015).

RESULTS

Obtaining Genome-Wide RNA Profiles for Serial Sections Along Antero-Posterior, Dorso-Ventral, and Left-Right Axes

In order to obtain a digital transcriptomic map of the catshark stage 17 forebrain, we excised three embryonic head explants

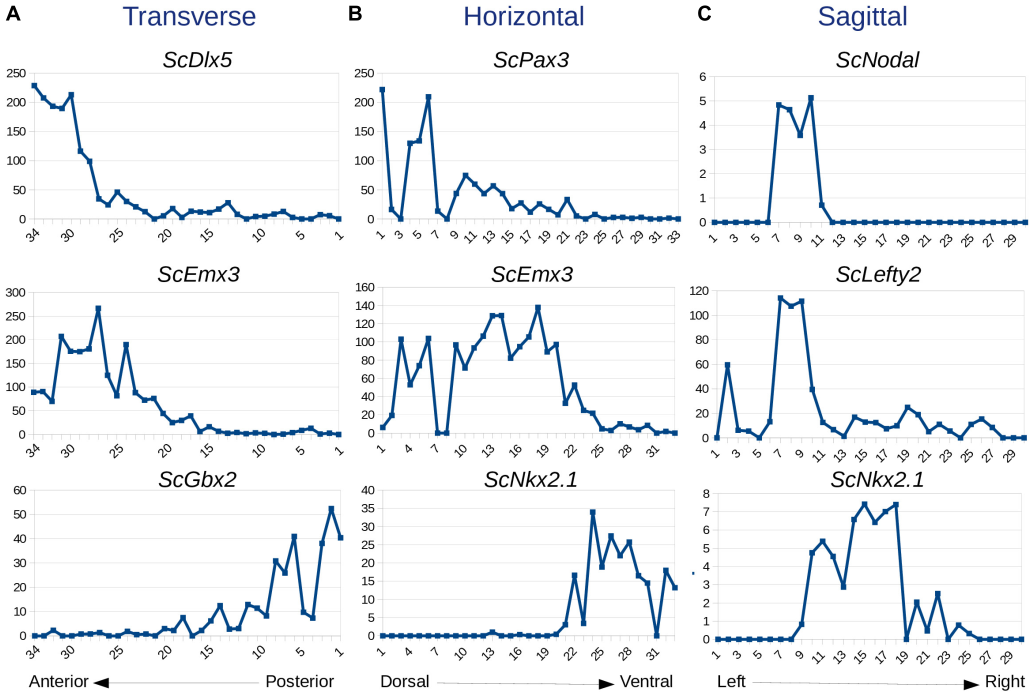


FIGURE 2 | Expression traces for selected genes along transverse **(A)**, horizontal **(B)**, and sagittal **(C)** planes. For each graph, section numbers and corresponding read counts are indicated along x- and y-axes, respectively. Sections 1–34, 1–33, and 1–30 are, respectively, numbered from posterior to anterior in **(A)**, dorsal to ventral in **(B)**, and left to right in **(C)**. Gene names are indicated at the top of each profile.

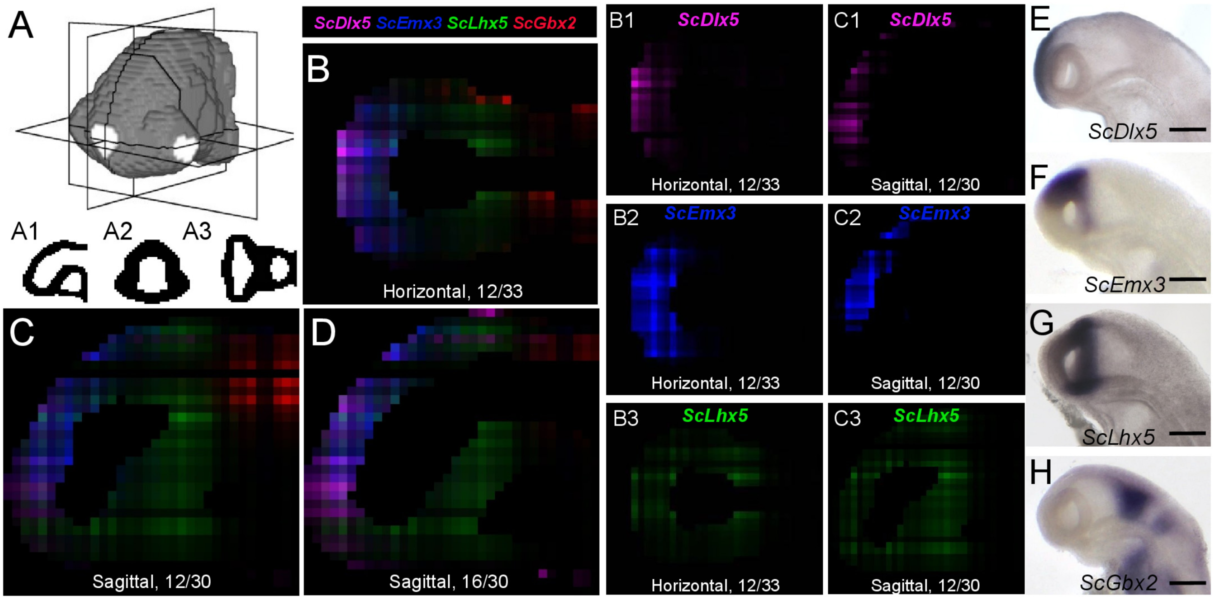

A
A1
A2
A3
ScDlx5 ScEmx3 ScLhx5 ScGbx2
B
Horizontal, 12/33
C
Sagittal, 12/30
D
Sagittal, 16/30
B1
ScDlx5
Horizontal, 12/33
C1
ScDlx5
Sagittal, 12/30
B2
ScEmx3
Horizontal, 12/33
C2
ScEmx3
Sagittal, 12/30
B3
ScLhx5
Horizontal, 12/33
C3
ScLhx5
Sagittal, 12/30
E
ScDlx5
F
ScEmx3
G
ScLhx5
H
ScGbx2

FIGURE 3 | Digital profiles reproduce regional patterns along AP axis. **(A)** Scheme representing the digitized model of stage 17 catshark head with sagittal, transverse and horizontal digital section planes shown as white rectangles. **(A1**,**A2**,**A3)** show examples of digitized sections along these planes. **(B–D)** Digital horizontal **(B)** and sagittal **(C,D)** sections showing merged profiles of *ScDlx5* (magenta), *ScEmx3* (blue), *ScLhx5* (green), and *ScGbx2* (red). Sections are numbered from dorsal to ventral in panel **(B)** and from left to right in panels **(C,D)**. **(B1–3**,**C1–3)** Digital sections showing the individual profiles, respectively, merged in panels **(B**,**C)** for *ScDlx5* **(B1,C1)**, *ScEmx3* **(B2,C2)**, and *ScLhx5* **(B3,C3)**. **(E–H)** Left lateral views of stage 17 catshark embryonic heads after ISH with probes for *ScDlx5* **(E)**, *ScEmx3* **(F)**, *ScLhx5* **(G)**, and *ScGbx2* **(H)**. Scale bars = 200 μm.

truncated anterior to the first branchial pouch (**Figure 1A**). Using cryostat sectioning along AP (antero-posterior), DV (dorso-ventral) and LR (left-right) axes, we obtained a total of 34 transverse, 33 horizontal, and 30 sagittal 18 mm frozen sections (**Figure 1B**). We next synthesized cDNA for each individual section using barcoded primers, which allowed to pool cDNA samples for linear amplification and quality controls of amplified products (**Figure 1C** and **Supplementary Table 3**). We next proceeded to Illumina library construction and sequencing to obtain about 210 million reads (PE-100). For the bioinformatic analysis, reads were demultiplexed and mapped onto a gene reference obtained taking advantage of the catshark genome NCBI annotation and of transcriptomic resources available in our laboratory (**Supplementary Tables 1, 2**), which led to the assignment of 76% reads to a reference contig (**Supplementary Figures 1A–C,D–F,G–I**). For each section plane, a minimum of 4 reads in at least two sections were detected prior to normalization for more than 13000 contigs (13354, 13501, and 14239 for sagittal, horizontal, and transverse planes, respectively). Read counts were then normalized, allowing the generation of expression profiles for any predicted gene along each section plane and of a genome-wide 3D digital model of gene expression (**Figures 1D,E**).

Expression Traces Along Antero-Posterior, Dorso-Ventral, and Left-Right Axes Reflect Expression Restrictions of Candidate Regional Markers

In order to test whether variations in read counts between serial sections along each axis reflect gene regional expressions, we focused on forebrain markers previously characterized in the catshark (*Dlx5*, *Emx3*, *Gbx2*, *Pax3*, *Nkx2.1*, *Nodal*, *Lefty2*; **Supplementary Table 4**). Markedly higher *Dlx5* and *Emx3* read counts are found in anterior, versus posterior, transverse sections in line with their telencephalic expression. Read counts drop to basal levels from sections 1–27 for *Dlx5* and from sections 1–20 for *Emx3*, which accurately reflects the more restricted anterior territory of the former (Derobert et al., 2002; Santos-Durán et al., 2015; **Figure 2A**). In contrast, read counts peak at posterior levels (sections 1–8) for *Gbx2*, a midbrain marker (**Figure 2A**). Along the DV axis, *Pax3* and *Emx3* read counts are high dorsally and return to basal levels in more ventral sections, with a more dorsal peak for the former, while *Nkx2.1* reads are only detected in ventral-most sections (22–33), in line with ISH profiles for these genes (Derobert et al., 2002; O’Neill et al., 2007; Santos-Durán et al., 2015; **Figure 2B**). Along the left-right axis, sharp left restricted peaks are observed for *Nodal* and *Lefty2* (sections 7–10), consistent with the asymmetric left diencephalic territory reported at this stage (Lagadec et al., 2015; **Figure 2C**). In contrast, the curve shows a plateau at the level of sections located on each side of the midline (sections 10–18) for the symmetrically expressed hypothalamic marker *Nkx2.1* (Santos-Durán et al., 2015; **Figure 2C**).

Construction of a 3D Digital RNA Profile Reproducing *in situ* Hybridization Expression Patterns

We next projected the sequence data obtained for each section onto a mask imaged by confocal microscopy (**Figure 1E**) to generate a 3D RNA profile. To do so, we used iterative proportional fitting (Fienberg, 1970), as described in Junker et al. (2014). This resulted in a digital 3D model, containing expression information computed by the algorithm, referred to hereafter as digital expression, for every gene of the reference database and in every voxel identified by its (x, y, z) coordinates in the mask. In order to test this model and assess whether virtual expression patterns reflect expected ISH expression profiles, we focused on forebrain regional markers known to exhibit conserved and highly specific expression patterns across vertebrates, either along AP (antero-posterior), DV (dorso-ventral), or LR (left-right) axis (**Figures 3–5**; **Supplementary Figure 2**; **Supplementary Table 4**). Along the AP axis, we visualized digital profiles for *ScDlx5*, *ScEmx3*, *ScLhx5*, and *ScGbx2* (**Supplementary Videos 1, 2**). *ScDlx5* and *ScEmx3* exhibit anterior digital expressions, while *ScLhx5* positive voxels are located in a more medial position, adjacent to a posterior *ScGbx2* territory, in line with ISH expressions (compare **Figures 3C,D,C1–C3** with **Figures 3E–H**). Similarly, as observed in ISH, *ScEmx3* digital territory extends further laterally and posteriorly than the one of *ScDlx5* (**Figures 3B,B1,B2**) and *ScLhx5* digital expression exhibits a sharp posterior border, contrasting with its anterior extension (**Figures 3B3,C3**). Similarly, along the DV axis, *ScNkx2.2* positive voxels are restricted to a medial and ventral location relative to optic evaginations, within a broader *ScSix3* digital territory (**Figures 4A,B,B2,B3** and **Supplementary Videos 3, 4**). The latter overlaps with *ScEmx3* digital expression, which extends further posteriorly and is restricted dorsally (**Figures 4A,B,B1,B2**). These features reproduce those observed in whole-mount ISH (**Figures 4C–E**). Along the left-right axis, left-restricted (*ScNodal*, *ScVg1*) or left-enriched (*ScLefty2*) digital signals are observed dorsally and posteriorly to optic evaginations, at the interface between *ScSix3* anterior and *ScIrx1l* posterior digital territories, as in whole-mount ISH (compare **Figure 5A–F** to **Figures 5G–K**; **Supplementary Videos 5–10**). Digital signals are primarily located dorsally as expected for *ScNodal*, *ScVg1*, and *ScLefty2* (**Figures 5A–C,G1–I1**), but a few additional positive voxels, without clear counterparts in ISH, can be observed more ventrally at the level of optic evaginations (**Figures 5A,C**). The accuracy of the digital profiles was further assessed by analysis of digital expressions of additional markers including paralogous genes (**Supplementary Figure 2** and **Supplementary Videos 11–16**). Comparison of *ScNkx2.1* and *Nkx2.2* digital profiles shows a broader expansion of the latter compared to its paralog along the AP axis, as observed in ISH (**Supplementary Figures 2A,A1,A2**). Along the same line, the sharp dorsal and ventral boundaries of *ScSix6* territory and its relative location within the broader *ScSix3* domain are reproduced by the digital profiles of the paralogs (compare **Supplementary Figures 2B,B1** to **Figures 2C,C1**), as well as the respective expression characteristics of *ScFgf17* (restricted to the anterior-most part of the forebrain)

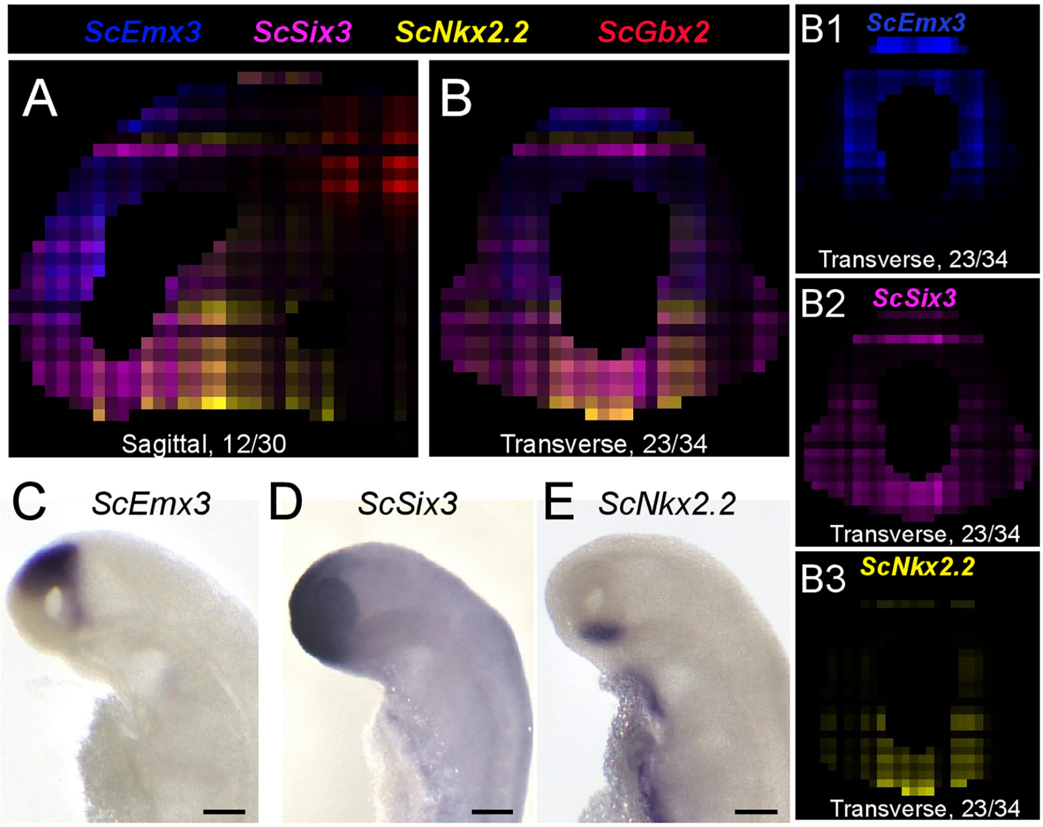


FIGURE 4 | Digital profiles reproduce regional patterns along DV axis. **(A,B)** Digital sagittal **(A)** and transverse **(B)** sections showing merged profiles of *ScEmx3* (blue), *ScSix3* (magenta), *ScNkx2.2* (yellow), and *ScGbx2* (red). Sections are numbered from left to right in panel **(A)** and from posterior to anterior in panel **(B)**. **(B1–3)** Digital sections showing the individual profiles merged in panel **(B)** for *ScEmx3* **(B1)**, *ScSix3* **(B2)**, and *ScNkx2.2* **(B3)**. **(C–E)** Left lateral views of stage 17 catshark embryonic heads after ISH with probes for *ScEmx3* **(C)**, *ScSix3* **(D)**, and *ScNkx2.2* **(E)**. Scale bars = 200 μm.

and *ScFgf8* (harboring a major midbrain expression and a more diffuse anterior signal) (**Supplementary Figure 2D**; compare **Supplementary Figures 2D1,2D2**). The ventral expansion of forebrain markers such as *ScFoxg1* or *ScFezF2* relative to *ScEmx3* is also correctly predicted by digital profiles (compare **Figure 3D** with **Supplementary Figures 2E,F,E1,E2,F1,F2**).

Autocorrelation as Estimate of the Spatial Restriction of Digital Profiles

In order to extract lists of genes exhibiting regionalized digital expression patterns, we computed spatial autocorrelations, commonly used to analyze the spatial structure of geographically distributed biological variables (Diniz-Filho et al., 2009) to the analysis of our digital expression model (**Figure 6** and **Supplementary Figure 3**). The underlying rationale is that genes displaying regionalized profiles should tend to exhibit similar digital expressions in neighboring voxels within their expression territory and therefore show high autocorrelation values (**Supplementary Figure 3A**). Moran's index, which provides a measure of spatial autocorrelation, and the *p*-value evaluating its statistical support, were calculated for all genes of the reference database (**Figure 6A** and **Supplementary Table 5**). In order to assess the relevance of this notion to the identification of regionalized genes, we first surveyed Moran's indexes (i) and *p*-values obtained for a selection of the forebrain regional markers analyzed above, exhibiting discrete digital territories of variable size (*ScFezF2*, *ScNkx2.2*, *ScEmx3*, *ScSix6*, *ScDlx5*, *ScFoxg1*, *ScVg1*, *ScIrx3*, *ScGbx2*, *ScFgf8*, *ScFgf17*, *ScNodal*: **Figure 6A**). In all cases, indexes > 0.1, supported by null *p*-values, were retrieved. The lowest index is obtained for *ScNodal* ($i = 0.11$), which has a relatively low total digital expression ($n = 619$ counts over all voxels), while highest indexes are obtained for *ScFezF2*, *ScScNkx2.2,* and *ScEmx3*, three genes exhibiting high digital expressions (total digital expression summed over all voxels > 14000) and a single, continuous territory. Reciprocally, we randomly selected genes within 5 sectors of the *p*-value versus Moran's index graph (sector 1: $0.01 < i < 0.02$; sector 2: $0.06 < -i < 0.07$; sector 3: $0.09 < i < 0.1$; sector 4: *p*-value = 0; and $0.1 < i < 0.2$: sector 5: *p*-value = 0 and $0.6 < i < 0.7$; **Figure 6A**). All genes within sector 4 and 5 exhibited strongly regionalized digital territories, with a digital expression observed in all voxels in the former and a majority of voxels in the latter (**Figures 6E,F** and **Supplementary Figures 4D,E**). In contrast, labeled voxels appeared more and more dispersed and generally less numerous from sector 3 to 1 (**Figures 6B–D** and

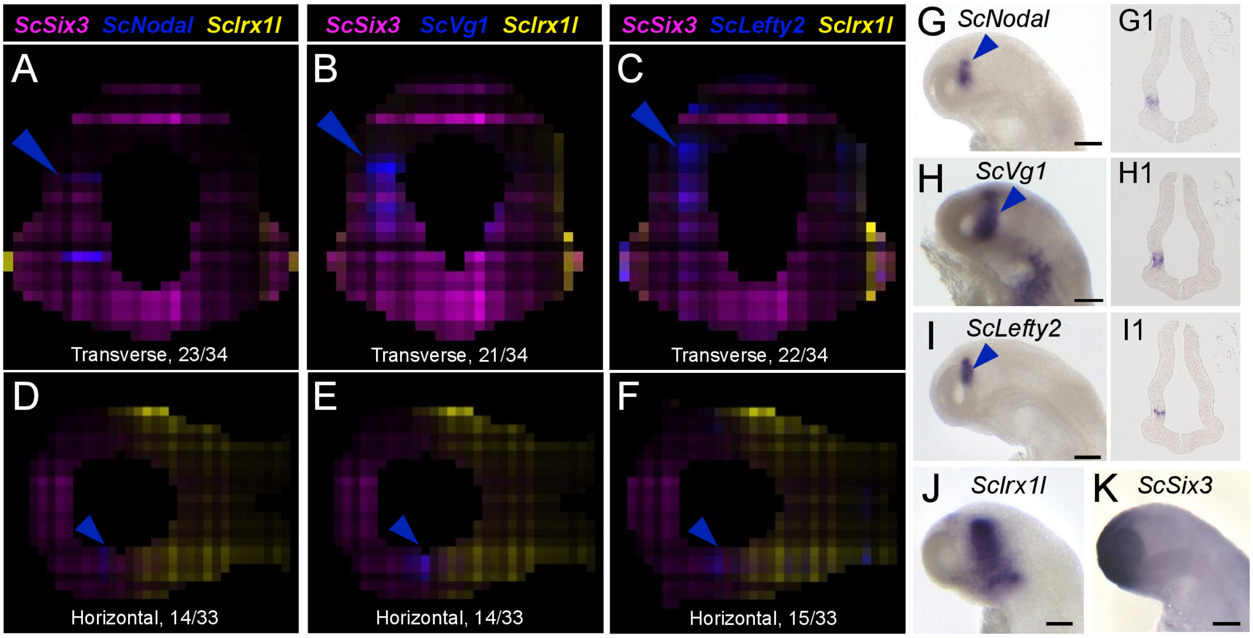


FIGURE 5 | Nodal signaling components exhibit left enriched profiles. **(A–F)** Digital transverse **(A–C)** and horizontal **(D–F)** sections showing merged profiles of *ScSix3* (magenta), *ScIrx1l* (yellow), and in blue *ScNodal* **(A,D)**, *ScVg1* **(B,E)**, *ScLefty2* **(C,F)**. Sections are numbered from posterior to anterior in panels **(A–C)** and from dorsal to ventral in panels **(D–F)**. **(G–K)** Left lateral views of stage 17 catshark embryonic heads after whole-mount ISH with probes for *ScNodal* **(G)**, *ScVg1* **(H)**, *ScLefty2* **(I)**, *ScIrx1l* **(J)**, and *ScSix3* **(K)**. **(G1–I1)** Horizontal sections of the brain of stage 17 catshark embryos after hybridization with probes for *ScNodal* **(G1)**, *ScVg1* **(H1)**, *ScLefty2* **(I1)**. Blue arrowheads point to *ScNodal*, *ScVg1*, and *ScLefty2* left-restricted digital or ISH signals in the dorsal diencephalon. Scale bars = 200 μm.

Supplementary Figures 4A–C). In general, genes of low total digital expression were over-represented among those with low autocorrelations and high *p*-values (**Supplementary Table 6**). For instance, more than 19% of the genes displaying –log(*p*-value) < 50 have a total digital expression < 100 and this value drops to, respectively, 2.1 and 0.1% for genes with $i < 0.1$ and $i > 0.1$, *p*-value = 0. Of note, 12100 genes, i.e., more than 60% of annotated genes, exhibit significant autocorrelation values ($i > 0.1$, *p*-value = 0; **Supplementary Table 5** and **Supplementary Figure 3B**).

A Correlation Statistical Approach Allows *de novo* Detection of Co-expressed Genes

The availability of a spatially resolved genome-wide profile allows the search for genes displaying specific expression criteria, for instance coexpression with selected subterritory markers. To exploit this possibility, we used a statistical approach relying on calculations of Pearson correlations (Cor) between pairs of genes (**Supplementary Figure 5A**). We first calculated Pearson correlations between digital expressions of all coding gene pairs taking over all voxels of our 3D model. Lists of genes ranked by correlation coefficient can then be extracted from this dataset for every single gene of the reference database. In order to assess the possibility to identify genes coexpressed with a given candidate by this method, we focused on correlations with *ScShh*, which shows a highly specific expression in the ventral midline, faithfully reproduced by the digital pattern, as proof-of-concept (**Figures 7A,B** and **Supplementary Table 7**). All 12 genes most strongly correlated to *ScShh* (Cor > 0.4) harbor a ventral digital expression, largely overlapping with the one of *ScShh* (**Figure 7A** and **Supplementary Figure 5B**). This list contains the catshark orthologs of two midline markers, respectively, coding for the Shh receptor Ptch1 and the extracellular matrix protein Slit3 (Ingham and McMahon, 2001; Yeo et al., 2001). In order to further validate the possibility of identifying novel genes coexpressed with *ScShh* by this approach, we conducted ISH for candidates retrieved from this list (*ScFoxa2*, *ScRrbp1*, *ScRop1l*, *ScIgfbp3*, *ScFoxb2*, *ScCcd39*). While no or very faint midline signals were observed for *ScRop1l* and *ScCcd39* (not shown), strong ventral restricted signals similar to *ScShh* were retrieved for *ScFoxa2,* as previously described (Ang and Rossant, 1994; Weinstein et al., 1994), but also for a set of genes not previously reported to display midline expression, such as *ScRrbp1*, *ScIgfbp3*, and *ScFoxb2* (**Figures 7B–F**). This validates both the predictive power of the 3D model and the possibility to extract novel information using a correlation approach.

DISCUSSION

Since the initial description of RNA tomography (Junker et al., 2014), this technique has been applied to a limited number of systems, including limbs, the regenerating or pathological heart and aorta formation, in major model organisms including

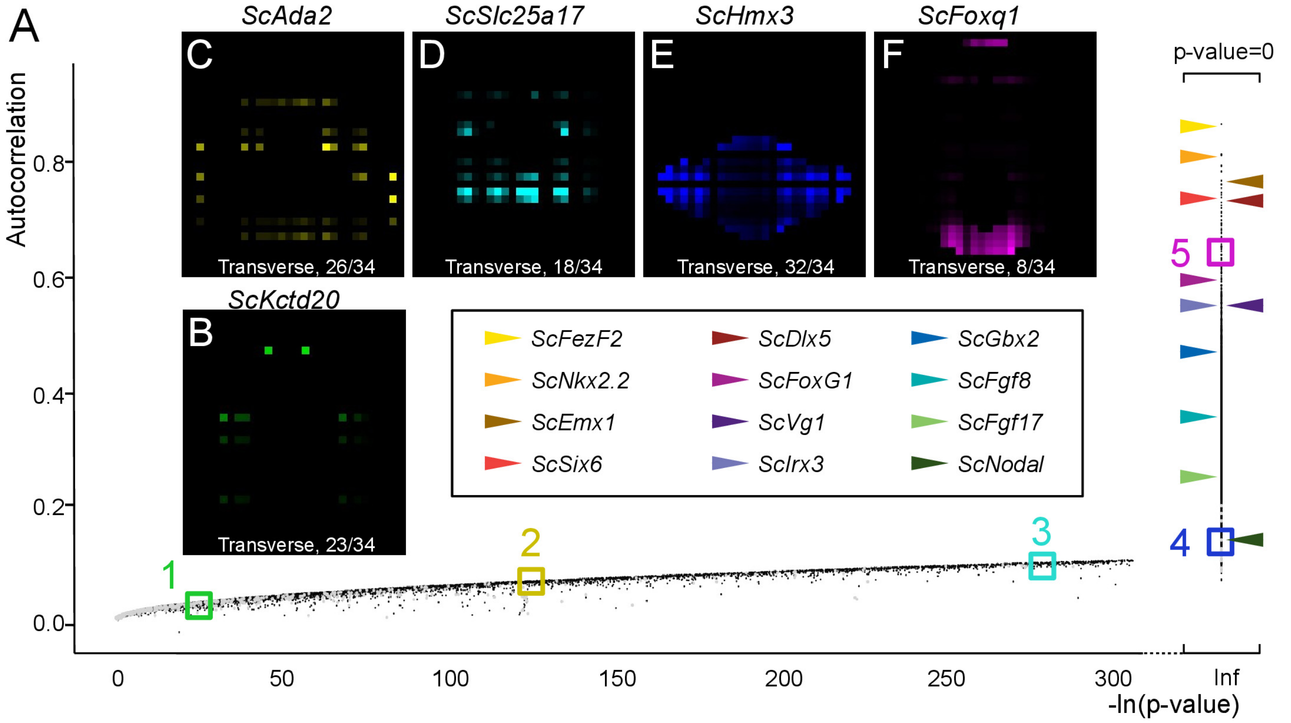


FIGURE 6 | Autocorrelation as an indicator of profile regionalization. **(A)** Graph showing autocorrelation values (y-axis) and their statistical support [-ln(*p*-value) on x-axis] for all gene models. Each dot corresponds to a single gene model, in gray if the corresponding total digital expression is below 100, otherwise in black. Gene models with null *p*-values are shown on the right of the graph with –ln(*p*-value) = infinite (Inf). Horizontal arrowheads point to dots corresponding to a selection of genes, exhibiting clearly regionalized expressions in ISH experiments. Color codes for these genes are shown boxed in black. Colored boxes delimit sectors of the graph numbered from 1 to 5 and containing gene models with decreasing *p*-values/increasing autocorrelations (green, sector 1; yellow, sector 2, cyan, sector 3; blue, sector 4; magenta, sector 5). **(B–F)** Digital transverse sections showing examples of profiles observed in sectors 1 **(B)**, 2 **(C)**, 3 **(D)**, 4 **(E)**, and 5 **(F)**. The corresponding gene identities are indicated and sections are numbered from posterior to anterior.

the zebrafish, mouse, chick, and *Drosophila* (Wu et al., 2016; Combs and Eisen, 2017; Burkhard and Bakkers, 2018; van den Brink et al., 2020; Yvernogeau et al., 2020; Holler et al., 2021). Most of these analyses have focused on characterizations along a single plane of interest, which is sufficient to identify genes differentially expressed along a polarity axis but does not provide a 3D spatial resolution. Here, we have applied this technique for the first time to a chondrichthyan, the catshark *S. canicula*, focusing on the developing head, shortly after neural tube closure. The resulting dataset is a digital model of this structure, containing expression information in each voxel of this volume for about 30,000 sequences, of which 21,000 correspond to coding ones. It thus provides a resource of interest to address forebrain regionalization, sensory organ, placode, or diencephalic asymmetry formation in the catshark and explore the conservation of these processes across jawed vertebrates.

In this analysis, we find that 3D digital expressions for known markers of broad forebrain subdivisions faithfully reproduce the broad characteristics observed by ISH, with the method being sensitive enough to allow the detection of regionalized expression for transcription factors or signaling molecules. Although they do not reach cellular resolution, digital profiles also correctly predict expression details, including differences between paralogs, such as *ScNkx2.1* and *ScNkx2.2*, or *ScSix3* and *ScSix6*, or the asymmetric left/dorsal restriction of *Lefty2*/*Nodal*/*Vg1* in the diencephalon. In principle, a lower limit in resolution in this approach is imposed by section thickness, here 18 mm. Albeit in the same order of magnitude, this exceeds cell diameters in catshark stage 17 embryos, which range from 5 to 15 mm. Decreasing section thickness to 10 mm might be a first approach to improve resolution. Since the Cell-Seq2 protocol used for library construction is widely employed for single cell sequencing, this should not result in a marked loss in sensitivity due to the lower RNA quantity per section, pending a possible increase of the number of linear amplification cycles. Another factor limiting resolution and inherent to tomo-seq is related to the 3D *post hoc* reconstruction. The approach indeed integrates data from three independent, non-identical biological samples, sectioned along non-perfect orthogonal planes, and involves a projection on an independently generated digital model of the tissue analyzed. Manual curation of this digital model, aimed at adjusting its shape and size to the 1D profiles of previously characterized candidate markers along each axis, may alleviate this difficulty. However, tomo-seq is unlikely to reach cellular resolution and as such, cannot be applied to the analysis of non-replicable biological samples such as tumors biopsies, for instance. A panel of alternative techniques, reaching high spatial resolution in high throughput gene expression analyses

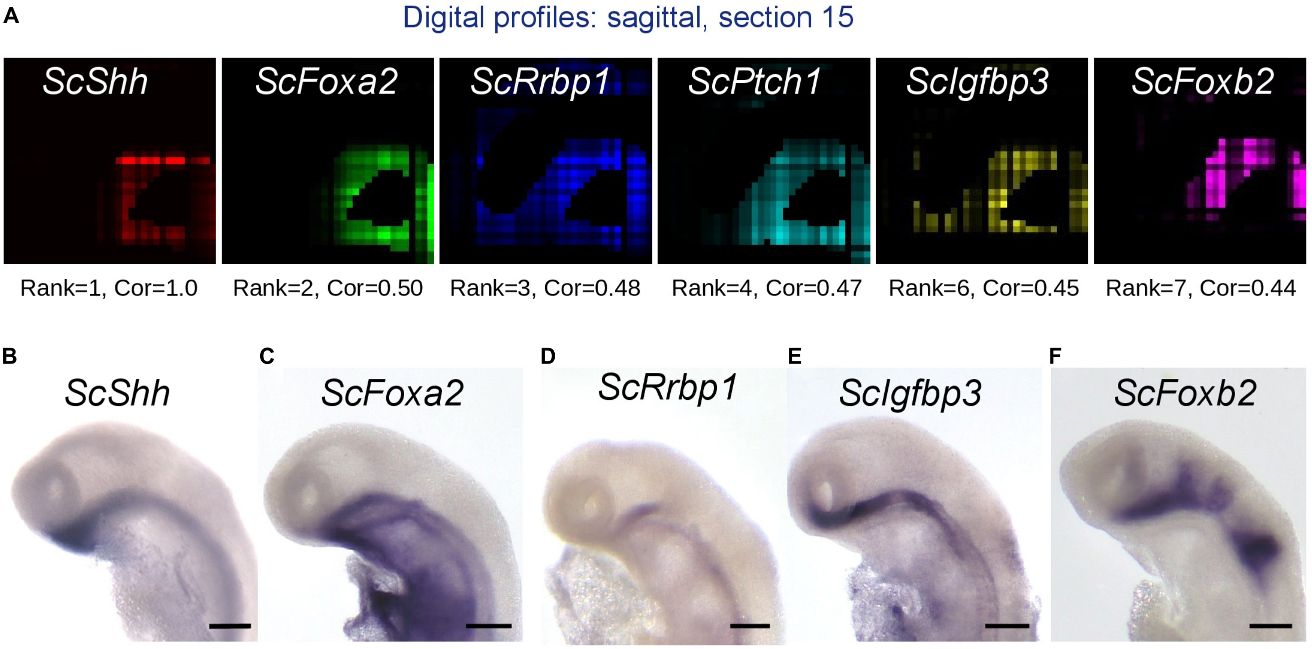


FIGURE 7 | Identification of genes harboring expression correlation with *ScShh*. **(A)** Digital profiles on sagittal section number 15 (out of 30, from left to right) of a selection of genes showing a high expression correlation to *ScShh* (red), *ScFoxA2* (green), *ScRrbp1* (blue), *ScPtch1* (cyan), *ScIgfbp3* (yellow), and *ScFoxb2* (purple). The rank of each gene in the list of gene models ordered by decreasing correlation to *ScShh* and the correlation value (Cor) are indicated below each section. **(B–F)** Left lateral views of stage 17 catshark embryonic heads after ISH using probes for *ScShh* **(B)**, *ScFoxA2* **(C)**, *ScRrbp1* **(D)**, *ScIgfbp3* **(E)**, and *ScFoxb2* **(F)**. Scale bars = 200 μm.

of non-replicable specimens, have emerged in the past few years. Multiplexed fluorescent *in situ* hybridization (FISH) protocols, such as hybridization chain reaction (HCR) or single molecule FISH (smFISH) based methods, allow *in situ* expression analyses of tens to hundreds of candidate genes with cellular and even subcellular resolution, and they have been successfully applied to a variety of species and tissues (Chen et al., 2015; Choi et al., 2016, 2018; Moffitt et al., 2018; Trivedi et al., 2018; Andrews et al., 2020). Other approaches, relying on *in situ* RNA capture from histological sections and *ex situ* sequencing, provide unbiased and genome wide characterizations, albeit with slightly lower spatial resolution (Ståhl et al., 2016; Asp et al., 2019; Rodriques et al., 2019; Vickovic et al., 2019; Ortiz et al., 2020). Compared to these technologies, tomo-seq presents a number of distinctive advantageous features including (1) a robust and fast protocol, without the need for lengthy imaging processes or experimental optimizations aimed, for instance, at monitoring the balance between *in situ* RNA accessibility and diffusion in the tissue, (2) the use of standard, inexpensive reagents and equipment for library construction, and (3) the production of genome-wide expression data, which can be screened rapidly for thousands of candidate genes, possibly following filtering on expression level and autocorrelation as described here. The combination of this technique with other approaches reaching cellular resolution, such as scRNA-seq, or FISH based methods applied to panels of selected genes, may be powerful approaches for comprehensive characterizations of tissue or organ RNA profiles (van den Brink et al., 2020).

An important perspective opened by the availability of this resource is the possibility to identify new genes characterized by a defined expression feature, such as co-expression with a territory marker. Established methodologies are still lacking for 3D tomo-seq datasets, which have been thus far only produced in the zebrafish and mouse. Junker et al. (2014) previously reported that a euclidean distance-based method could lead to the detection of new genes displaying expression similar to an organizer marker. Here we used an alternative statistical method, relying on estimates of Pearson correlations between any pair of genes and over all voxels, to address this question. This approach retrieves genes known or expected to be co-expressed with *Shh* such as *Foxa2*, *Ptch1*, or *Slit3*, with high correlation values. It also identifies genes, such as *Rrbp1* or *Igfbp3*, that have to our knowledge not previously been reported as co-expressed with *Shh*. We also validate some of these profiles by ISH, which demonstrates the potential of this approach to explore the entire repertoire of genes expressed in territories of interest. Such exhaustive characterization is important to identify synexpression groups of genes possibly submitted to the same regulatory framework, but also to assess territory homologies based on a significant number of signature markers in comparative approaches.

Similarly, we used spatial autocorrelation as a measure to identify genes exhibiting regionalized expression. The underlying mathematical notion has been widely used in biology (Diniz-Filho et al., 2009), albeit not to tomo-seq data. Its adequacy in this context is related to the assumption that expression values per voxel are non-independent variables, neighboring voxels, for instance around a peak or trough, tending to behave similarly. This characteristic is indeed observed for most regionalized genes, albeit not for those exhibiting salt-and-pepper expression patterns, such as Notch targets (Caprioli et al., 2002). In line with the relevance of this indicator, strongly supported positive autocorrelation values (p-value ∼ 0) were obtained for all highly regionalized genes analyzed, whether expressed in broad territories like most of the forebrain markers tested, or in highly restricted domains such as left-restricted *Vg1* or *Nodal*. Reciprocally, randomly selected patterns exhibit an expression dispersal per voxel, which increases inversely proportionally to their autocorrelation value. These data point to autocorrelation as a relevant indicator to identify genes of regionalized expression. Its use suggests that as previously proposed (Junker et al., 2014), regionalized genes (p-value = 0, autocorrelation > 0.1) represent an important fraction of the total gene repertoire.

In summary, our results show that tomo-seq is an effective method to characterize broad partitions of organs or embryonic structures. The recurrent need for such analyses in the evo-devo field does not necessarily require cellular resolution. One limitation, however, is that the size of the tissue of interest be sufficient for a section-based approach: tomo-seq would, for instance, be inadequate for a similar forebrain characterization shortly after neural tube closure in the zebrafish, due to the small size of its embryonic head at this stage. More generally, large body size is an obvious limitation for the maintenance of large populations under laboratory conditions. This characteristic has thus been counter-selected in the choice of major model organisms in developmental biology, such as *Drosophila*, zebrafish, or *C. elegans*. However, large size becomes an advantage for section-based spatial transcriptomics techniques such as tomo-seq. These techniques provide novel tools in evo-devo models harboring large embryo and body sizes such as the catshark, which although non-amenable to genetic approaches, has become a reference for questions pertaining to the origin and ancestral characteristics of jawed vertebrates.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI (accession: PRJNA758756).

ETHICS STATEMENT

Ethical review and approval was not required for the animal study because the work only used non-mammalian embryos, at a stage preceding central nervous system differentiation. This study does not require approval from an animal ethics committee according to national and EU regulations.

AUTHOR CONTRIBUTIONS

RL, AQ, ML, AM, and LM conducted the experimental work (sectioning, Illumina library construction, imaging, and ISH). HM, KM, and SD took charge in bioinformatics (sequence processing and analysis, 3D reconstruction) and statistical analyses. SM, RL, and PB conceived the work, analyzed the data, and wrote the manuscript with inputs from all authors. All authors had full access to all data, approved the article, and took responsibility for the integrity of the data and the accuracy of the data analysis.

FUNDING

This work was funded by Agence Nationale de la Recherche contract n° ANR-16-CE13-0013-02 to SM and PB and Ph.D. fellowships to ML (contract LSP n°156393 Région Occitanie) and LM (Ministère de la Recherche, ED515).

ACKNOWLEDGMENTS

We thank Pascal Romans, the Aquariology Service of the Banyuls Oceanological Observatory (OOB) and Michael Fuentes for providing and maintaining catsharks, EMBRC-France for support to local marine infrastructures, David Pecqueur and the BioPic imaging platform for access to confocal microscopy, Nyree West and the Bio2Mar service for access to molecular biology platform, the UMR7232 Service de Bio-informatique BSBII for HM's support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.744982/full#supplementary-material

REFERENCES

Andrews, T. G. R., Gattoni, G., Busby, L., Schwimmer, M. A., and Benito-Gutiérrez, È (2020). Hybridization chain reaction for quantitative and multiplex imaging of gene expression in amphioxus embryos and adult tissues. *Methods Mol. Biol.* 2148, 179–194. doi: 10.1007/978-1-0716-0623-0_11

Ang, S. L., and Rossant, J. (1994). HNF-3β is essential for node and notochord formation in mouse development. *Cell* 78, 561–574. doi: 10.1016/0092-8674(94)90522-3

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., et al. (2019). A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell* 179, 1647–1660. doi: 10.1016/j.cell.2019.11.025

Ballard, W. W., Mellinger, J., and Lechenault, H. (1993). A series of normal stages for development of *Scyliorhinus canicula*, the lesser spotted dogfish *Chondrichthyes*: *Scyliorhinidae*. *J. Exp. Zool.* 267, 318–336. doi: 10.1002/jez.1402670309

Burkhard, S. B., and Bakkers, J. (2018). Spatially resolved RNA-sequencing of the embryonic heart identifies a role for Wnt/β-catenin signaling in autonomic control of heart rate. *ELife* 7:e31515. doi: 10.7554/eLife.31515

Caprioli, A., Goitsuka, R., Pouget, C., Dunon, D., and Jaffredo, T. (2002). Expression of notch genes and their ligands during gastrulation in the chicken embryo. *Mech. Dev.* 116, 161–164. doi: 10.1016/S0925-4773(02)00136-3

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348:aaa6090. doi: 10.1126/science.aaa6090

Choi, H. M. T., Calvert, C. R., Husain, N., Huss, D., Barsi, J. C., Deverman, B. E., et al. (2016). Mapping a multiplexed zoo of mRNA expression. *Development* 143, 3632–3637. doi: 10.1242/dev.140137

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., et al. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* 145:dev165753. doi: 10.1242/dev.165753

Combs, P. A., and Eisen, M. B. (2017). Genome-wide measurement of spatial expression in patterning mutants of *Drosophila melanogaster*. *F1000Res.* 6:41. doi: 10.12688/f1000research.9720.1

Coolen, M., Menuet, A., Chassoux, D., Compagnucci, C., Henry, S., Lévèque, L., et al. (2008). The dogfish *Scyliorhinus canicula*: a reference in jawed vertebrates. *Cold Spring Harb. Protoc.* 2008, 1–14. doi: 10.1101/pdb.emo111

Derobert, Y., Plouhinec, J.-L., Sauka-Spengler, T., Le Mentec, C., Baratte, B., Jaillard, D., et al. (2002). Structure and expression of three Emx genes in the dogfish *Scyliorhinus canicula*: functional and evolutionary implications. *Dev. Biol.* 247, 390–404. doi: 10.1006/dbio.2002.0700

Diniz-Filho, J. A. F., Nabout, J. C., de Campos Telles, M. P., Soares, T. N., and Rangel, T. F. L. V. B. (2009). A review of techniques for spatial modeling in geographical, conservation and landscape genetics. *Genet. Mol. Biol.* 32, 203–211. doi: 10.1590/S1415-47572009000200001

Fienberg, S. E. (1970). An iterative procedure for estimation in contingency tables. *Anna. Math. Stat.* 41, 907–917. doi: 10.1214/aoms/1177696968

Holler, K., and Junker, J. P. (2019). RNA tomography for spatially resolved transcriptomics (tomo-seq). *Methods Mol. Biol.* 1920, 129–141. doi: 10.1007/978-1-4939-9009-2_9

Holler, K., Neuschulz, A., Drewe-Boß, P., Mintcheva, J., Spanjaard, B., Arsiè, R., et al. (2021). Spatio-temporal mRNA tracking in the early zebrafish embryo. *Nat. Commun.* 12:3358. doi: 10.1038/s41467-021-23834-1

Ingham, P. W., and McMahon, A. P. (2001). Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* 15, 3059–3087. doi: 10.1101/gad.938601

Junker, J. P., Noël, E. S., Guryev, V., Peterson, K. A., Shah, G., Huisken, J., et al. (2014). Genome-wide RNA tomography in the zebrafish embryo. *Cell* 159, 662–675. doi: 10.1016/j.cell.2014.09.038

Lagadec, R., Laguerre, L., Menuet, A., Amara, A., Rocancourt, C., Péricard, P., et al. (2015). The ancestral role of nodal signalling in breaking L/R symmetry in the vertebrate forebrain. *Nat. Commun.* 6:6686. doi: 10.1038/ncomms7686

Lee, J. H. (2017). De novo gene expression reconstruction in space. *Trends Mol. Med.* 23, 583–593. doi: 10.1016/j.molmed.2017.05.004

Moffitt, J. R., Bambah-Mukku, D., Eichhorn, S. W., Vaughn, E., Shekhar, K., Perez, J. D., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* 362, eaau5324. doi: 10.1126/science.aau5324

O'Neill, P., McCole, R. B., and Baker, C. V. H. (2007). A molecular analysis of neurogenic placode and cranial sensory ganglion development in the shark, *Scyliorhinus canicula*. *Dev. Biol.* 304, 156–181. doi: 10.1016/j.ydbio.2006.12.029

Ortiz, C., Navarro, J. F., Jurek, A., Märtin, A., Lundeberg, J., and Meletis, K. (2020). Molecular atlas of the adult mouse brain. *Sci. Adv.* 6:eabb3446. doi: 10.1126/sciadv.abb3446

Plouhinec, J. L., Leconte, L., Sauka-Spengler, T., Bovolenta, P., Mazan, S., and Saule, S. (2005). Comparative analysis of gnathostome Otx gene expression patterns in the developing eye: implications for the functional evolution of the multigene family. *Dev. Biol.* 278, 560–575. doi: 10.1016/j.ydbio.2004.11.019

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., et al. (2019). Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 363, 1463–1467. doi: 10.1126/science.aaw1219

Santos-Durán, G. N., Menuet, A., Lagadec, R., Mayeur, H., Ferreiro-Galve, S., Mazan, S., et al. (2015). Prosomeric organization of the hypothalamus in an elasmobranch, the catshark *Scyliorhinus canicula*. *Front. Neuroanat.* 09:37. doi: 10.3389/fnana.2015.00037

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82. doi: 10.1126/science.aaf2403

Sun, S., Zhu, J., and Zhou, X. (2020). Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* 17, 193–200. doi: 10.1038/s41592-019-0701-7

Trivedi, V., Choi, H. M. T., Fraser, S. E., and Pierce, N. A. (2018). Multidimensional quantitative analysis of mRNA expression within intact vertebrate embryos. *Development* 145:dev156869. doi: 10.1242/dev.156869

van den Brink, S. C., Alemany, A., van Batenburg, V., Moris, N., Blotenburg, M., Vivié, J., et al. (2020). Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* 582, 405–409. doi: 10.1038/s41586-020-2024-3

Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., et al. (2019). High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* 16, 987–990. doi: 10.1038/s41592-019-0548-y

Waylen, L. N., Nim, H. T., Martelotto, L. G., and Ramialison, M. (2020). From whole-mount to single-cell spatial assessment of gene expression in 3D. *Commun. Biol.* 3:602. doi: 10.1038/s42003-020-01341-1

Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M., et al. (1994). The winged-helix transcription factor HNF-3β is required for notochord development in the mouse embryo. *Cell* 78, 575–588. doi: 10.1016/0092-8674(94)90523-1

Wu, C. C., Kruse, F., Vasudevarao, M. D., Junker, J. P., Zebrowski, D. C., Fischer, K., et al. (2016). Spatially resolved genome-wide transcriptional profiling identifies BMP signaling as essential regulator of zebrafish cardiomyocyte Regeneration. *Dev. Cell* 36, 36–49. doi: 10.1016/j.devcel.2015.12.010

Yeo, S. Y., Little, M. H., Yamada, T., Miyashita, T., Halloran, M. C., Kuwada, J. Y., et al. (2001). Overexpression of a slit homologue impairs convergent extension of the mesoderm and causes cyclopia in embryonic zebrafish. *Dev. Biol.* 230, 1–17. doi: 10.1006/dbio.2000.0105

Yvernogeau, L., Klaus, A., Maas, J., Morin-Poulard, I., Weijts, B., Schulte-Merker, S., et al. (2020). Multispecies RNA tomography reveals regulators of hematopoietic stem cell birth in the embryonic aorta. *Blood* 136, 831–844. doi: 10.1182/blood.2019004446

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Mayeur, Lanoizelet, Quillien, Menuet, Michel, Martin, Dejean, Blader, Mazan and Lagadec. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 29 October 2021
doi: 10.3389/fcell.2021.749806

JNK Mediates Differentiation, Cell Polarity and Apoptosis During Amphioxus Development by Regulating Actin Cytoskeleton Dynamics and ERK Signalling

Ildiko M. L. Somorjai[1,2,3]*, Matthias T. Ehebauer[4†], Hector Escrivà[2] and Jordi Garcia-Fernàndez[3,5]

[1] School of Biology, University of St Andrews, St Andrews, United Kingdom, [2] Sorbonne Université, CNRS, Biologie Intégrative des Organismes Marins, Observatoire Océanologique, Banyuls-sur-Mer, France, [3] Departament de Genètica, Microbiologia i Estadística, University of Barcelona, Barcelona, Spain, [4] European Molecular Biology Laboratory, Hamburg Unit, Hamburg, Germany, [5] Institut de Biomedicina, University of Barcelona, Barcelona, Spain

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Hiroshi Wada,
University of Tsukuba, Japan
Michel Vervoort,
UMR 7592 Institut Jacques Monod (IJM), France

***Correspondence:**
Ildiko M. L. Somorjai
imls@st-andrews.ac.uk

†Present address:
Matthias T. Ehebauer,
Modcor, Didcot, United Kingdom

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: 30 July 2021
Accepted: 17 September 2021
Published: 29 October 2021

Citation:
Somorjai IML, Ehebauer MT, Escrivà H and Garcia-Fernàndez J (2021) JNK Mediates Differentiation, Cell Polarity and Apoptosis During Amphioxus Development by Regulating Actin Cytoskeleton Dynamics and ERK Signalling. Front. Cell Dev. Biol. 9:749806. doi: 10.3389/fcell.2021.749806

c-Jun N-terminal kinase (JNK) is a multi-functional protein involved in a diverse array of context-dependent processes, including apoptosis, cell cycle regulation, adhesion, and differentiation. It is integral to several signalling cascades, notably downstream of non-canonical Wnt and mitogen activated protein kinase (MAPK) signalling pathways. As such, it is a key regulator of cellular behaviour and patterning during embryonic development across the animal kingdom. The cephalochordate amphioxus is an invertebrate chordate model system straddling the invertebrate to vertebrate transition and is thus ideally suited for comparative studies of morphogenesis. However, next to nothing is known about JNK signalling or cellular processes in this lineage. Pharmacological inhibition of JNK signalling using SP600125 during embryonic development arrests gastrula invagination and causes convergence extension-like defects in axial elongation, particularly of the notochord. Pharynx formation and anterior oral mesoderm derivatives like the preoral pit are also affected. This is accompanied by tissue-specific transcriptional changes, including reduced expression of *six3/6* and *wnt2* in the notochord, and ectopic *wnt11* in neurulating embryos treated at late gastrula stages. Cellular delamination results in accumulation of cells in the gut cavity and a dorsal fin-like protrusion, followed by secondary Caspase-3-mediated apoptosis of polarity-deficient cells, a phenotype only partly rescued by co-culture with the pan-Caspase inhibitor Z-VAD-fmk. Ectopic activation of extracellular signal regulated kinase (ERK) signalling in the neighbours of extruded notochord and neural cells, possibly due to altered adhesive and tensile properties, as well as defects in cellular migration, may explain some phenotypes caused by JNK inhibition. Overall, this study supports conserved functions of JNK signalling in mediating the complex balance between cell survival, apoptosis, differentiation, and cell fate specification during cephalochordate morphogenesis.

Keywords: JNK, apoptosis, amphioxus, chordate, ERK, cellular extrusion, Wnt, notochord

INTRODUCTION

Embryogenesis requires the accurate temporal and spatial coordination of tissue fate, cell movements, and cell numbers in a controlled and reproducible way. In most metazoans, this is achieved by the complex orchestration of only seven evolutionarily conserved developmental pathways, including canonical Wnt, JAK/STAT, TGFb, Hh, Nuclear Receptor, Notch/Delta, and receptor tyrosine kinase (RTK) signalling (Babonis and Martindale, 2017). Small changes in this equilibrium can either result in a complete failure of development, with the generation of "monsters," or, if increasing fitness over evolutionary time, in the diversity of organismal form and function seen today.

Downstream of these extracellular signals, mitogen activated protein kinase (MAPK) signalling is an ancient mechanism of signal transduction that has diversified significantly within the major eukaryote lineages (Xu et al., 2017; Kalapos et al., 2019), and whose dysregulation can lead to a number of diseases, including cancers (Kim and Choi, 2010). Three conventional MAPK subfamilies have been described: the extracellular signal regulated kinase (ERK), the p38 MAP kinase (p38), and the c-Jun N-terminal kinase (JNK) pathways (Krens et al., 2006; Lavoie et al., 2020). RTKs such as FGFR and EFGR mediate some of their cellular responses through the phosphorylation and activation of ERK (Shilo, 2014). ERK is often considered to be a survival factor, but it also plays important roles in cell behaviour, differentiation, and fate (Lavoie et al., 2020). JNK, on the other hand, tends to be associated with cellular migration and polarity, often as part of a non-canonical (β-catenin-independent) Wnt pathway (Gao and Chen, 2010). JNK also regulates apoptosis and the cell cycle (Pinal et al., 2019), and, similarly to p38 (Canovas and Nebreda, 2021), some stress responses.

During development, JNK proteins play a crucial role in epithelial fusion events, for instance dorsal or thorax closure in *Drosophila* (Martín-Blanco et al., 2000; Ríos-Barrera and Riesgo-Escovar, 2013), and neural tube closure in vertebrates such as mouse and chick (Krens et al., 2006; Pai et al., 2012). JNK signalling is also involved in diverse processes including gastrulation in echinoderms (Long et al., 2015); convergence extension (CE) movements in *Xenopus* (Yamanaka et al., 2002; Kim and Han, 2005); neuronal polarisation (Castro-Torres et al., 2020), spinal cord neuron development and axonal pathfinding (Schellino et al., 2019), and lateral line neuromast hair cell development and regeneration in zebrafish (Cai et al., 2016; He et al., 2016); and tail regression during metamorphosis in ascidians (Chambon et al., 2007), to only list a few. Thus, JNK is a multifunctional protein with a complex range of tissue-specific and temporally regulated activities in developing vertebrate and invertebrate embryos.

To dissect the evolution of the JNK signalling cascade in shaping morphogenetic trajectories requires the identification of both its conserved and derived functions using a comparative approach outside conventional model systems. Cephalochordates (also known as lancelets or amphioxus) represent a key phylogenetic branch for understanding the evolution of developmental mechanisms at the invertebrate to vertebrate transition (Bertrand and Escrivà, 2011). They share many fundamental developmental processes with vertebrates, including regulative development, neurulation, and somitogenesis, as well as key chordate characters such as a notochord, hollow nerve cord, post-anal tail, and the endostyle, a thyroid hormone producing organ (Bertrand and Escrivà, 2011). However, broadly speaking, lancelet genomes are much simpler, with often single orthologues of genes found in multiple copies in vertebrates, which have arisen due to the whole genome duplications in this lineage (Escrivà, 2018). They also lack the complex acquired immune systems and behaviours characteristic of many model systems (Yuan et al., 2014).

Recent research in amphioxus has provided considerable knowledge on the mechanisms underlying axial patterning, germline formation, nervous system development, regeneration, genome architecture, and gene regulation (Somorjai et al., 2012a,b; Marlétaz et al., 2018; Lin et al., 2020; Bozzo et al., 2021). Much of this has been facilitated by recent advances in "omics," including transcriptomics, genomics and single cell, and the advancement of new methodologies including knockdown and microscopy (Escrivà, 2018). At this time, the community has achieved a good understanding of the gene repertoires of major signalling pathways, such as RTKs (D'Aniello et al., 2008), or how core signalling pathways such as Wnt (Somorjai et al., 2018; Kozmikova and Kozmik, 2020), Nodal (Li et al., 2017), and FGF (Bertrand et al., 2011), among others, govern patterning and cell fate. However, comparatively little is still known about the cellular mechanisms downstream underlying some of the key processes governing morphogenesis.

Here, we aimed to fill this gap in our understanding by studying the role of JNK signalling in regulating amphioxus development. An early study identified a single JNK gene in the genome of the Florida amphioxus, *Branchiostoma floridae* (Bertrand et al., 2009). Moreover, JNK is known to undergo Nova-mediated alternative splicing in amphioxus, similarly to vertebrates (similar direction of exon inclusion events, Irimia et al., 2011). However, nothing has so far been reported on a role for JNK signalling in morphogenesis. Based on research in other systems, we hypothesised that amphioxus JNK might regulate cellular behaviours during gastrulation and neurulation, such as epithelial fusion, neural tube closure and convergence-extension processes. Here, we show that pharmacological inhibition of JNK activity using the small molecule SP600125 causes dramatic shortening of the A/P axis and considerable cellular disruption, including loss of many anterior structures and perturbed notochord formation. We propose that an appropriate balance of JNK signalling is required to maintain cellular integrity and polarity, and that mis-regulation results in cells delaminating through changes in the actin cytoskeleton and phosphorylated ERK (pERK), culminating in Caspase-mediated cell death of depolarised cells. We also show that at least some of these phenotypes are accompanied by changes in transcriptional timing, loss, and expansion of gene expression. Specifically, we see early upregulation of *wnt11* in posterior ectoderm concomitant with reduced expression in somites and at the hinge, and an almost complete loss of *six3/6* and *wnt2* in the notochord. Taken together we provide a plausible scenario for how JNK activity orchestrates complex morphogenetic processes to govern axial elongation and differentiation to ultimately shape the cephalochordate body plan.

MATERIALS AND METHODS

Embryo Collection and Treatments

Adult *Branchiostoma lanceolatum* were collected in Argelès-sur-Mer, France, by manual sieving of substrate as previously described (Fuentes et al., 2007). Spawning was according to previously published methods (Fuentes et al., 2007). Briefly, gravid males and females were subjected to a thermal shock for 36 h, returned to ambient temperature (19°C) and allowed to spawn naturally after sunset in individual cups containing seawater. Sperm and eggs were mixed in 0.22 μM filtered seawater to ensure consistent and timed fertilisation and maintained at 19°C for the entire developmental period.

Drug treatments and control vehicle were administered directly into petri dishes containing filtered seawater and embryos at appropriate stages, in this case right at the blastula stage (B stage; staging according to Carvalho et al., 2021) and at the end of gastrulation (G6), when the neural plate could be seen to begin flattening but while the blastopore was still open. For JNK inhibition, the cell permeable c-Jun N-terminal kinase inhibitor SP600125 (JNK Inhibitor II, Calbiochem) was used at concentrations of 2.5, 5, or 10 μM after initial pilot experiments. For Caspase inhibition, the cell-permeable, irreversible pan-Caspase Caspase inhibitor I (Z-VAD-fmk, Calbiochem) was used at 200 μM. DMSO was added at an equivalent concentration for all control experiments. Embryos were fixed in freshly prepared 4% PFA in MOPS buffer in filtered seawater overnight at 4°C, followed by several washes and storage in 70% ETOH at −20°C (WMISH), or for 2 h at room temperature (RT) followed by at least three washes and storage in 1× NaPBS at 4°C (immunohistochemistry and Phalloidin staining, respectively). For SP600125 and Z-VAD-fmk and combined inhibitor titration experiments, 50–100 embryos were used at each concentration. Definitive experiments were performed on 2–3 independent batches of embryos derived from single fertilisations (i.e., gametes from one mother and father) in four spawning seasons (2008–2011), each consisting of 500–1000 embryos each, to guarantee reproducibility. This resulted in batches of at least 200 embryos per treatment group and time point.

Immunohistochemistry

Treated embryos and stage-matched controls were prepared for immunohistochemistry and F-Actin staining as per previous protocols (Somorjai et al., 2012a,b) unless otherwise described. Briefly, after three washes in 1× TBS (Tris Buffered Saline), embryos were permeabilised using 0.5–1% Triton in 1× TBS for 40 min at RT. After copious washing in 1× TBS supplemented with 0.1% Tween (TBST), embryos were incubated with primary antibodies overnight with rocking at 4°C. After multiple washes in 1× TBST, a blocking step was included for 2 h at RT in 5% heat inactivated sheep serum supplemented with 2% BSA prior to adding secondary antibodies. To detect Caspase-3-mediated cell death, the anti-cleaved Caspase-3 (cCasp-3) primary antibody (Asp 175; Cell Signaling Technology) was used at 50 μM. Monoclonal antibody clones DP311 and DP312 (kindly provided by Nipam Patel, University of Berkeley, CA, United States) were used at 1:30 as previously described (Somorjai et al., 2012b) and label both Pax3/7 protein products (Barton-Owen et al., 2018). Proliferation was monitored using the phospho-histone H3 (PH3) antibody at 1:500 (Millipore). α-Acetylated tubulin (Sigma) was used at 1:500 to label cilia and axons, the C2206 antibody (Sigma) to label β-catenin at the membranes (Somorjai et al., 2012a) and the phosphorylated p44/42 MAPK (ERK1/2) Thr202/Tyr204 antibody (Cell Signaling Technology) to label pERK. When used, F-Actin was visualised using Alexa Fluor 488 or Alexa Fluor 568 conjugated Phalloidin (1:200, Molecular Probes) at RT for 2–4 h or in the final overnight washes along with 1:400 Alexa Fluor secondary antibodies (Molecular Probes) at 4°C. Embryos were rinsed in 1× TBST, then washed in DAPI staining solution (1:5000 of 5 mg/ml stock) for 20–40 min at RT to visualise nuclei, followed by final washes and mounting in ProLong Gold Antifade medium for fixed cells (Thermo Fisher Scientific). All immunohistochemistry experiments were conducted at least in duplicate on 5–10 embryos per control/treatment group. A random selection of embryos was mounted on slides for imaging by confocal microscopy (3–10 embryos). Confocal images were acquired with a Leica SPII confocal microscope. Images were post-processed in Fiji ImajeJ version 2.1.0/1.53c (Schindelin et al., 2012) and Adobe Photoshop 2021.

Whole Mount *in situ* Hybridisation

Wnt1, *wnt2*, *wnt3*, *wnt4*, *wnt5*, *wnt6*, *wnt7*, *wnt8*, *wnt9*, and *wnt11* probes are from Somorjai et al. (2008) and Somorjai et al. (2018); *pax3/7*, *chordin*, *brachyury2*, *neurogenin*, and *MRF1* (originally *MyoD*) are from Somorjai et al. (2008); *otx*, *foxQ2*, and *six3/6* are from Albuixech-Crespo et al. (2017); *musashi* is from Dailey et al. (2016); and *snail* is from Bertrand et al. (2011). Fragments of *dkk1/2/4*, *dkk3*, *sfrp1/2/5*, *sfrp3/4*, *fz1/2/7*, *fz4*, *fz5/8*, and *fz9/10* genes were amplified by PCR using a "touchdown" programme (Don et al., 1991) on pooled embryonic cDNA and cloned into pGEMT-Easy (Promega) or pBlueScript II KS+ (Invitrogen) vectors using standard procedures (Dailey et al., 2016). See **Supplementary File 1** for sequences of probes used. WMISH was performed as previously described (Somorjai et al., 2008) after stage-dependent permeabilization with 1.125 U/ml Proteinase K (Roche). DIG-labelled (Roche) antisense probes were *in vitro* transcribed using T3, SP6, or T7 polymerase (NEB) as appropriate. The chromogenic reaction was performed in the dark for 2 h or up to several days, with daily changes of staining solution, using BMPurple (Roche) or NBT/BCIP (Roche). Embryos were mounted on slides in 80% glycerol and imaged using a Leitz DMRB microscope (Leica Microsystems) with Nomarski optics. All WMISH experiments were conducted at least in triplicate with 10–20 embryos per control/treatment group. Consistency of expression was assessed, and a representative selection was photographed (3–10 embryos). Images were post-processed in Adobe Photoshop 2021.

Bioinformatics and Comparative Structure Modelling

Branchiostoma floridae and *Branchiostoma belcheri* JNK sequences were identified by tBLASTn search using human MAPK8, MAPK9, and MAPK10 protein sequences using the publicly available NCBI database.[1] The *B. lanceolatum* gene model BL14252 was similarly identified by querying the European lancelet genome on EnsemblMetazoa using BLAST with the human sequences.[2] However, as the gene model appeared to have an unlikely deletion, we used transcriptomic data to manually curate the sequence to generate an almost complete consensus sequence (see **Supplementary Figure 1**).

A comparative model of the *B. belcheri* JNK MAP kinase, residues N47 to S403, was created using Modeller 10.1 (Webb and Sali, 2016) based on the coordinates of residue D45 to N400 of human MAPK10 (1PMV) (Scapin et al., 2003). A structure-based sequence alignment of the amphioxus and human sequence was created using Clustal Omega[3] and manually edited to ensure that secondary structural elements remained gap-free. The sequence identity was 84%. Model geometry was validated using PROCHECK (Laskowski et al., 1993). The Ramachandran plot showed that the comparative model had 94.9% of residues in the most favoured conformations and 5.1% in the additionally allowed regions. There were no residues with disallowed conformations. The superposition of 1PMV and the model gives a root mean square deviation of 0.42 Å for the backbone atoms.

RESULTS

JNK Plays a Role in Invagination and Elongation in Amphioxus

In *Drosophila*, there is a single JNK orthologue, but in humans, there are three JNK genes (JNK1, 2, and 3 or MAPK 8, 9, and 10) resulting from the two whole genome duplications in the ancestor of the vertebrate lineage (Dehal and Boore, 2005). A previous study identified a single JNK orthologue in the *B. floridae* genome (Bertrand et al., 2009). However, because the increase in genomic resources available for amphioxus species has highlighted a number of unexpected lineage specific genome duplications (e.g., Barton-Owen et al., 2018), we took advantage of more recent data, including availability of the European amphioxus (*B. lanceolatum*) genome (Marlétaz et al., 2018), to validate this result. BLAST search using the human proteins confirmed the existence of a single orthologue of JNK in *B. lanceolatum*, *B. belcheri*, and *B. floridae* (**Supplementary Figure 1**). Sequence similarity suggests that they are highly conserved, and structurally similar to vertebrate proteins. We therefore surmised that a pharmacological approach using the small molecule inhibitor SP600125 (Bennett et al., 2001), which has affinity for all three vertebrate orthologues, would specifically target amphioxus JNK function.

Comparative modelling of the structure of amphioxus JNK bound to SP600125 relative to known conformations of MAPK8 and MAPK10 with the inhibitor revealed that both the binding pocket and the drug's docking sites are conserved (**Figure 1A**). Taken together, these data suggest that a pharmacological approach using SP600125 to inhibit JNK activity will be highly specific in amphioxus.

We initially tested several SP600125 concentrations from 1 to 20 μM, beginning at blastula or late gastrula stages. All embryos treated up to 5 μM at blastula stages (B stage) arrested during gastrulation (227/227; compare to DMSO controls 0/247; two-tailed $z = 11.74$, $P < 0.00001$), producing cap-shaped or exogastrulae (**Supplementary Figure 2**, and see below), but were alive based upon the normal though delayed initiation of ciliary movement. Those treated at late gastrula stages (G6) consistently developed shortened axes relative to controls (**Figure 1B**), with the strength of phenotype occurring in a dose-dependent manner from 2.5 to 10 μM. All embryos, regardless of treatment concentration, continued to swim in a directional spiralling fashion toward light sources (not shown). At the stronger concentrations and later stages of development, besides being short and stubby, embryos possessed a dorsal fin-like protrusion (cyan arrowhead), leading to the designation of "*orca*" phenotype. Over time, cells also apparently accumulated within the archenteron or gut cavity (black arrowheads), suggesting that JNK inhibition leads to polarity or CE defects through cellular extrusion into luminal spaces.

Gene Expression in Anterior Endoderm and the Notochord Are Modulated by JNK Activity

Analysis at the gross morphological level suggested that, though smaller, the *orca* embryos nevertheless possessed all major tissue types, including nerve cord, notochord, and somites, albeit disorganised. Embryos allowed to develop to late neurula (N4) or pre-mouth stages (N5 or T0) further appeared to lack anterior mesodermal/endodermal structures. In order to assess more specifically which tissues were affected, we performed WMISH using a bank of 28 genes, including markers for the ectoderm and nervous system (e.g., *neurogenin*, *distalless*, and *snail*), somites and muscle (e.g., *pax3/7* and *MRF1*), the notochord (*brachyury*, *chordin*, *musashi*, and *six3/6*), anterior endoderm (*otx* and *six3/6*) as well as markers for embryonic polarity (e.g., *nodal*, *foxQ2*). We also included a large number of Wnt signalling components, from ligands (*wnt1–9* and *wnt11*) to receptors (*fz1/2/7*, *fz4*, *fz5/8*, and *fz 9/10*) and antagonists and co-receptors (*sfrp4*, *sfrp1/2/5*, *dkk1/2/4*, and *dkk3*) for two reasons. First, the Wnt pathway defines a wide range of cell and tissue types in amphioxus (Somorjai et al., 2018). Second, there is some evidence, mostly in vertebrates, for cross-talk between the JNK and other Wnt pathways, both at protein and transcriptional levels (Hardy et al., 2008; Lee et al., 2009; Seo et al., 2010; Mazzotta et al., 2016). JNK has even been shown to phosphorylate β-catenin to regulate adherens junctions and cellular adhesion (Lee et al., 2009; Zeke et al., 2016). We therefore hypothesised that (1) convergence-extension defects might be uncovered through the analysis of markers for neural

---

[1] https://blast.ncbi.nlm.nih.gov/

[2] https://metazoa.ensembl.org/

[3] https://www.ebi.ac.uk/Tools/msa/clustalo/

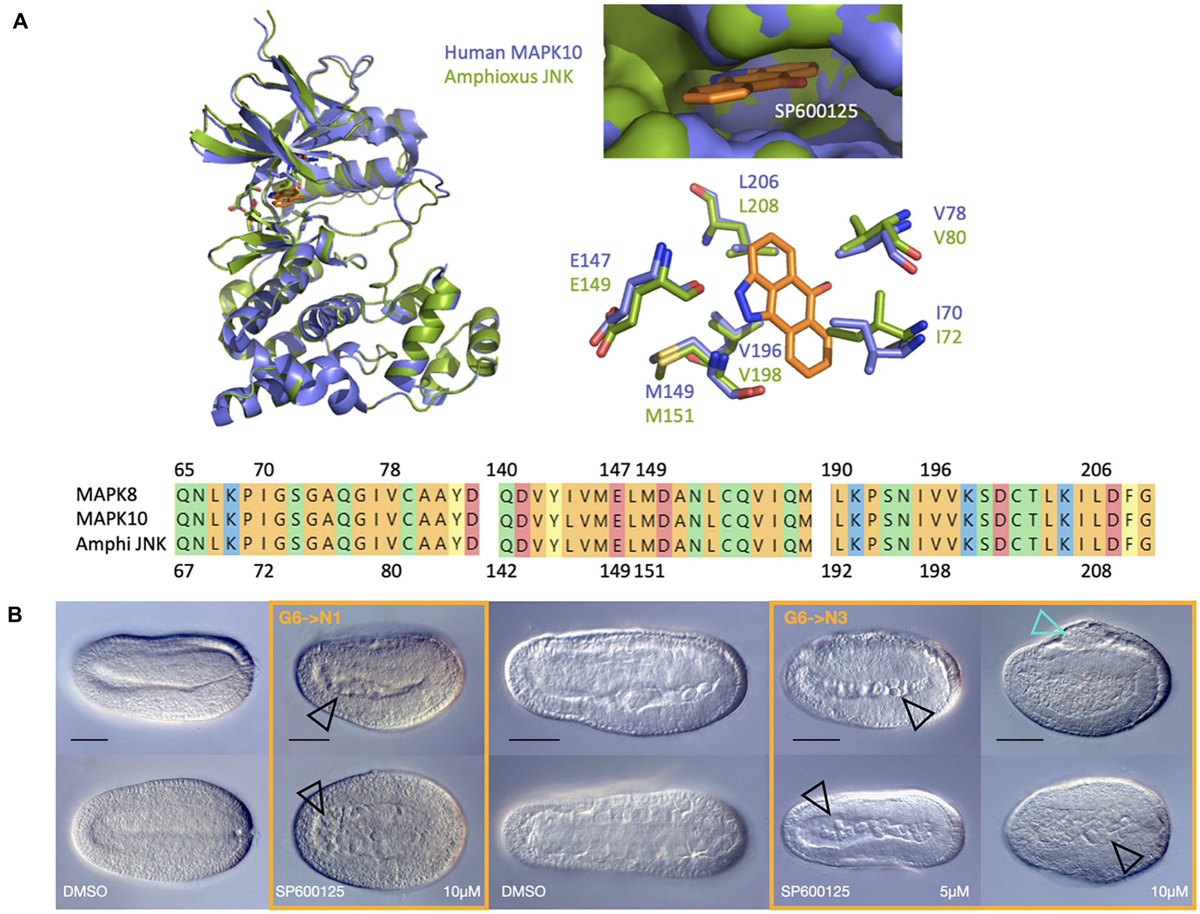

A
Human MAPK10
Amphioxus JNK
SP600125
L206
L208
V78
V80
E147
E149
I70
I72
V196
V198
M149
M151
65 70 78 140 147 149 190 196 206
MAPK8 QNLKPIGSGAQGIVCAAYD QDVYIVMELMDANLCQVIQM LKPSNIVVKSDCTLKILDFG
MAPK10 QNLKPIGSGAQGIVCAAYD QDVYLVMELMDANLCQVIQM LKPSNIVVKSDCTLKILDFG
Amphi JNK QNLKPIGSGAQGIVCAAYD QDVYLVMELMDANLCQVIQM LKPSNIVVKSDCTLKILDFG
67 72 80 142 149 151 192 198 208
B
G6->N1
G6->N3
DMSO
SP600125
10μM
DMSO
SP600125
5μM
10μM

FIGURE 1 | SP6000125 is a specific JNK inhibitor in amphioxus and causes axis elongation defects. **(A)** Comparative modelling of the amphioxus JNK (green) in complex with the small molecule inhibitor SP600125 (orange) using human MAPK10 as a template (purple) reveals conservation of docking sites at conserved residues (top). The alignment of human MAPK8, human MAPK10, and amphioxus JNK shows 100% identity of residues within the actives site (bottom).
(B) SP6000125 causes dose-dependent phenotypes when amphioxus embryos are treated at the late gastrula stage. The top row shows lateral views, with anterior to the left and dorsal up. The bottom row shows dorsal views. Treatments are boxed in orange, with their stage matched DMSO controls to the left. Black outlined arrowheads highlight an accumulation of cells within the gut after only short treatments with 10 μM SP600125. Over time, all embryos exposed to the JNK inhibitor show elongation defects, and by N3, in addition to accumulation of cellular material within the archenteron (black arrowheads), treated embryos have fin-like dorsal protrusions (cyan outlined arrowhead) at 10 μM SP600125. Representative embryos are shown. DMSO control, N1 *n* = 170; 10 μM SP600125, N1 *n* = 126; 5 μM SP600125, N3 *n* = 158. Scale bars = 50 μM.

and notochord tissues, specifically, although morphogenesis of other tissues may also be affected; and (2) we might detect transcriptional and protein-level responses of the Wnt pathway, including feedback loops, to treatment with SP600125.

We focused primarily for this analysis on embryos treated with 10 μM and allowed to progress to mid-neurula stages (N2 or N3) as this produced a good balance between strength of phenotype and good progression of development and differentiation. Gene expression assessed in embryos treated at the blastula stages was either too difficult to compare with early neurula stage controls due to the early arrest in gastrulation, or closely matched that seen in wild type gastrulae (**Supplementary Figure 2**). We did not assess these early phenotypes further here.

Taking into consideration the fact that embryos were generally more compact, and therefore some compression or distortion of expression domains is to be expected, our candidate approach revealed four classes of result. In some cases, a particular marker may show no expression change for one anatomical structure or germ layer, and yet differ significantly between controls and treated embryos for another, and thus may find itself within more than one group.

Genes Showing Broad Maintenance of Normal Gene Expression

Many genes showed no major change in gene expression. Generally, markers for embryonic polarity fall within this category. This includes *wnt1*, expressed at the most posterior of the embryo in the tailbud; *fz5/8*, expressed in anterior ectoderm and endoderm; and *nodal*, expressed asymmetrically on the left side in both treated and control embryos (**Figure 2A**). *Brachyury2*, *chordin*, and *musashi* appear to be expressed normally within the notochord domains, although the domain itself may be somewhat mediolaterally expanded in line with the shortened axis. *Pax3/7*, *wnt8*, and *MRF1* are expressed in somites. Neural domains are broadly congruent, as assessed by *chordin*, *pax3/7*, and *neurogenin*, although the latter shows clear disorganisation of neural positioning (**Figure 2A**, yellow asterisks), possibly reflecting epithelial fusion defects caused by JNK inhibition. *Dll* is expressed most strongly in anterior and posterior ectoderm, with some expression along neural folds, in both control and treated embryos. Other genes exhibiting no discernible change in expression domain or inconclusive results include *fz1/2/7*, *fz4*, *fz9/10*, *wnt7*, *wnt6*, *wnt3*, *wnt4*, and *sfrp1/2/5*, at least when assessed by N2/N3 (**Supplementary Figure 3** and not shown).

Subtle Changes in Extent of Expression or Patterning Within Specific Domains

We also found that some expression domains either extended further than expected or were reduced relative to the controls. For instance, anterior ventral endodermal expression of *dkk3* was expanded dorsally in treated embryos to include cells in close apposition to, or possibly within the anterior neural plate/cerebral vesicle (**Figure 2A**, black outlined arrowhead). Conversely, *six3/6* expression was maintained in the cerebral vesicle and anterior-most endoderm but was absent in presumptive anterior notochord (as per Albuixech-Crespo et al., 2017) in treated embryos (white outlined arrowhead). Likewise, the dorsal-ventral extent of expression of *foxQ2* appeared similar in control and treated embryos, but in fact extended further in lateral domains when viewed dorsally (black arrowhead). Anterior expression of *snail* was also reduced in treated embryos relative to controls. Finally, *dkk1/2/4* appeared strongly expressed specifically in the anterior-most pair of somites as well as the anterior neural plate/cerebral vesicle (black outlined arrowhead) in treated embryos compared to controls.

Ectopic Expression

Generally, true ectopic expression was rare and isolated to only few cells and may reflect tissue disorganisation resulting from JNK inhibition. In treated embryos, we found cases of clusters of cells outside their normal domain expressing *dll* and *wnt1* (**Figure 2A**, black outlined arrowheads) in ectodermal neural fold and tailbud cells, respectively. More convincingly, in treated embryos allowed to develop to the N3 stage, we also saw strong ectopic expression of *wnt11* in the posterior ectoderm of the tail fin (black outlined arrowhead). This domain of expression appears much later in wild type embryos, after N4. We also found that treatment with SP600125 resulted in the cerebral vesicle and anterior endoderm expression domains of *otx* to merge, reflecting either ectopic expression or a loss of separation between the two tissue types (**Supplementary Figure 3**, black outlined arrowhead).

Loss of Expression

Very few genes demonstrated a complete loss of expression, with three exceptions. *Dkk3* shows expression within a cluster of ventral endodermal cells in control embryos, a pattern consistently absent in treated embryos (**Figure 2A**, white asterisk). However, we found that this phenotype had variable penetrance even in wild type embryos (not shown), possibly reflecting slight delays in development, and should be considered with caution. *Wnt11* expression appears highly reduced in the last enterocoelic somites after treatment at N2, despite strong posterior ventral endoderm expression, and appears absent in the presumptive notochord cells within the hinge region of the tailbud (**Figure 2B**, white outlined arrowhead). Allowing embryos to develop further until at least N3 results in expression patterns resembling those of much later embryos (wild type N4 stage, rightmost panel), with the exception of continued absence of hinge expression (white outlined arrowhead), as well as complete absence of anterior ventral endoderm expression (yellow outlined arrowhead), which is characteristic of matched N3 controls but not of later N4 stages (filled yellow arrowhead). This suggests a loss of expression or cells within the anterior endodermal *wnt11* domain. Most notable in this category is expression of *wnt2* in treated embryos at N3, which have strong expression in at least two (sometimes three) clusters within the nerve cord, similar to stage-matched control embryos (**Figure 2C**). In contrast, they entirely lack notochordal *wnt2* expression (white outlined arrowhead), or anterior endodermal expression (outlined yellow arrowheads) relative to the control (filled yellow arrowhead). Although also lacking endodermal expression, the earlier N2 stage wild type embryos have consistent notochordal expression. Extension of the chromogenic reaction only strengthened the nerve expression, with background staining in other tissues. This result is particularly striking when compared to the strong expression of notochord markers, in particular *brachyury2*.

Taken together, and as expected if one of the functions of JNK is to mediate cellular behaviour, our results indicate that transcriptional changes in many key patterning and cell type specification genes are relatively unaffected by JNK inhibition both during gastrulation and neurulation. However, we did see several marked changes, particularly of gene expression domains in the anterior endoderm, as well as shifts in relative contribution to different tissues of genes such as *wnt11*, and most notably, loss of *six3/6* and *wnt2* in the notochord.

Inhibition of JNK Causes Significant Cellular Disruption

Given the *orca* phenotype consists of a shortening of the axis, some tissue disorganisation and a number of changes in gene expression domains, but that broadly speaking cell and tissue types appear to be maintained, we wanted to assess how JNK inhibition might be mediating these effects at the cellular level. We therefore utilised Phalloidin staining to assess any changes in F-Actin as a readout for modulation of the cytoskeleton and cellular behaviours, as many cytoskeletal proteins are substrates of JNK, including Actin-binding proteins and β-catenin (Lee et al., 2009; Zeke et al., 2016). In combination with other antibodies, including against β-catenin, Pax3/7, and acetylated-tubulin, we identified several tissue-specific changes that might

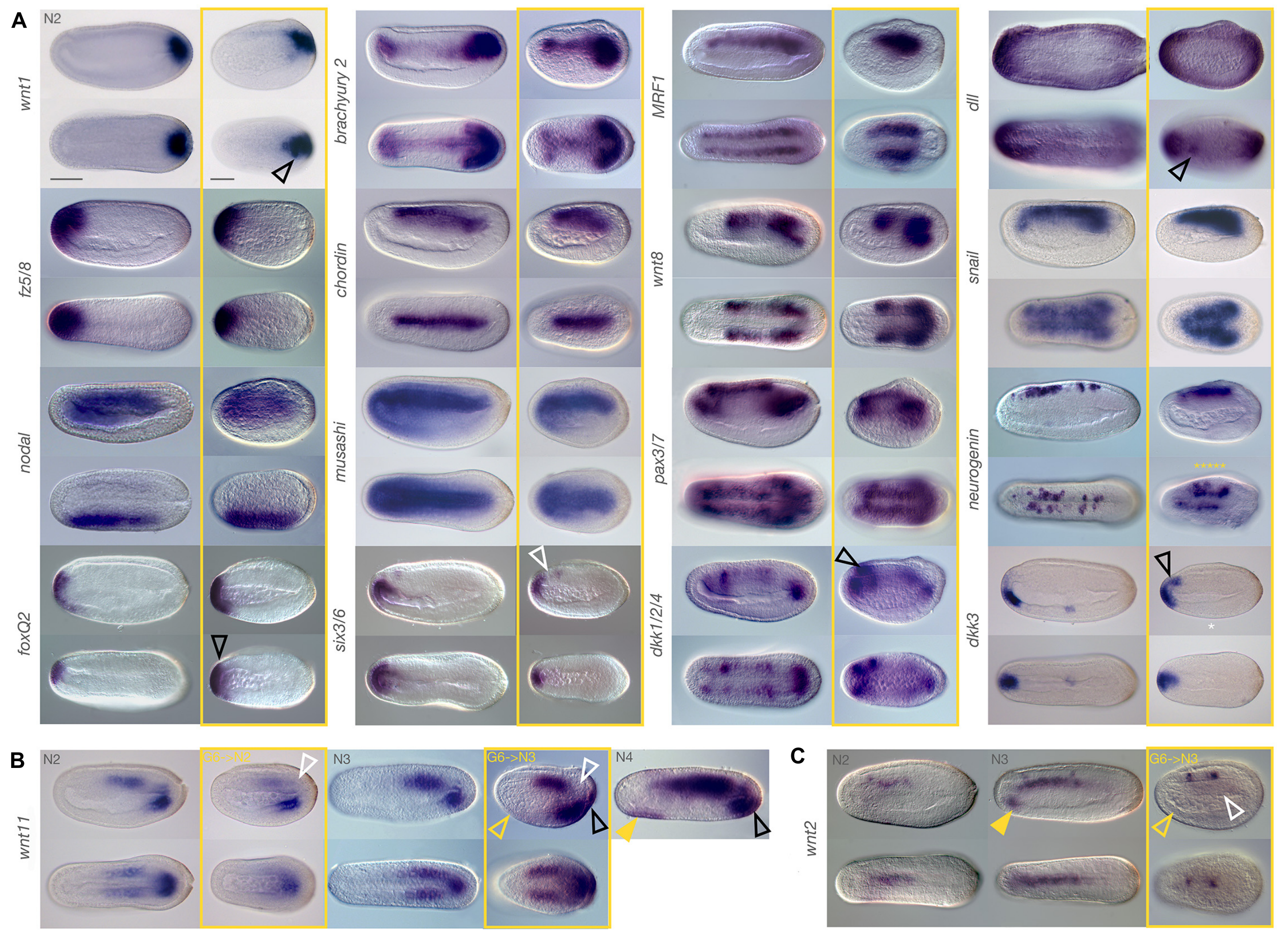


FIGURE 2 | Marker gene expression after JNK inhibition at late gastrula stages is modulated. Embryos were treated with 10 μM SP600125 and allowed to develop to N2, N3, or N4 stages (boxed in yellow). Stage matched controls are on the left. **(A)** Treated embryos generally show similar expression patterns to the controls. Exceptions include *dkk1/2/4*, *dkk3*, *dll*, *foxQ2*, and *wnt1* which show slightly expanded or ectopic expression (black outlined arrowheads), and *six3/6*, which appears to lack expression in the anterior notochord (white outlined arrowhead) despite comparable neural expression. Irregular expression of *neurogenin* in neurons along either side of the midline (yellow asterisks) is also apparent. **(B)** After treatment with the JNK inhibitor, *wnt11* expression is reduced relative to control N2 embryos (white arrowhead). By N3, chordoneural hinge expression is still absent (white outlined arrowhead), and there is no ventral endoderm expression (yellow outlined arrowhead) similarly to controls. However, ectopic expression within posterior ectoderm of the tailfin is observed (black outlined arrowhead). This is more similar to N4 embryos (black outlined arrowhead) than to the stage matched N3 controls, but does not fit with anterior ventral endoderm expression in these later stage embryos (filled yellow arrowhead). **(C)** Expression of *wnt2* is also dysregulated by JNK inhibition in the notochord (white outlined arrowhead) and in anterior ventral endoderm (yellow outlined arrowhead), relative to stage matched N3 controls, which also have expression in the notochord (compare to N2 embros). This occurs in spite of normal neural expression, and cannot be explained by a systemic delay of development. Representative embryos are shown. Scale bars = 50 μM.

contribute to the effects of JNK inhibition on morphogenesis in amphioxus embryos.

Disorganisation of the CNS

One of the most striking observations of the gene expression analysis was the apparent asymmetry of neuron alignment at N3 stages seen with *neurogenin* staining. We hypothesised that defects in early stages of neurulation, either of fusion of the epithelial sheets or of neural folds, might account for this patterning defect. At N1, when neurulation has begun with fusion of the epithelial cells along the dorsal midline, we saw no major differences in F-Actin between treated and control animals. In both, F-Actin accumulates at the leading edge of the zippering epithelial sheet (**Figures 3A,A′**), and by N3 the epidermis overlying the neural plate/tube is continuous (**Figures 4A,A′,C,C′**).

We also analysed the distribution of β-catenin, the nuclear effector of canonical Wnt signalling and a key component of the actin cytoskeleton at adherens junctions and membranes linking F-Actin to E-Cadherin. We found no obvious changes in accumulation of β-catenin at the adherens junctions or the membranes at early or later stages (**Figures 3B,C,C′**), and neurulation appears to reach completion. We did, however, see that the cells accumulating within the gut cavity of SP600125-treated embryos show central aggregation of F-Actin, which was not matched

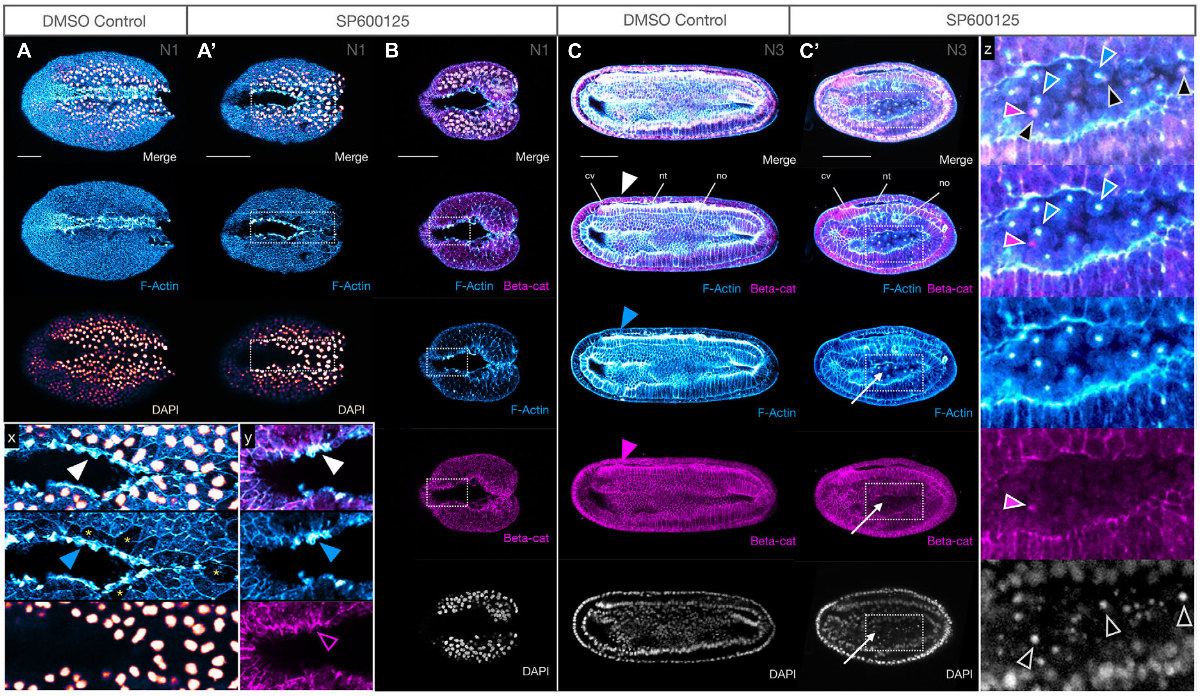


FIGURE 3 | JNK inhibition produces no major defects in epidermal fusion or neural tube closure. **(A,A′)** Phalloidin staining of F-Actin (blue) in apical membranes shows epidermal fusion in SP600125-treated embryos proceeds similarly, but more slowly, in treated embryos **(A′)** relative to DMSO controls **(A)** at N1 stage. Inset x clearly shows an accumulation of F-Actin in the leading edge (white arrowheads) of fusing epidermal sheets along the midline. Cells with such protrusions experiencing tension after having just fused, or in the process of fusion, appear to be depleted in cytoplasmic F-Actin (yellow asterisks). Dorsal views are shown. **(B)** Immunohistochemistry of β-catenin (beta-cat, magenta) shows that the accumulation of F-Actin in leading edge cells is not accompanied by concomitant accumulation of β-catenin (inset y, arrowheads). Dorsal views shown. **(C,C′)** By N3, the neural folds have fused into a closed neural tube in both controls **(C)** and SP600125-treated embryos **(C′**, arrowheads) despite obvious morphogenesis defects. Cells accumulating within the archenteron contain central aggregates of F-Actin (white arrows) and pycnotic nuclei (DAPI) (inset z, coloured arrowheads show non-correspondence of nuclear fragments and F-Actin aggregates, and general absence of β-catenin except as infrequent and generally microscopic puncta). Representative embryos are shown. Lateral views shown. In all cases, anterior is to the left. cv, cerebral vesicle; nt, neural tube; no, notochord. Scale bars = 50 μM.

by changes in β-catenin expression (**Figure 3C′**, inset z) and appear to be dying as assessed by the presence of pycnotic nuclei.

Next, we assessed the organisation of the neuronal cells along the neural folds using DP311 and DP312 clones, which have previously been shown to label Pax3/7 (Somorjai et al., 2012b). In control N2 embryos, we found expression of Pax3/7 in two stripes along the dorsal midline, corresponding to neurons within the neural folds (**Figures 4A,A′,A″,B**). JNK inhibition results in a delay of neural plate folding, but Pax3/7 is expressed at the edges as expected for an earlier stage of development (**Figure 4C′**, red arrows), and by N3 expression is broadly as expected along fused neural folds (**Figure 4D**, red arrows). However, upon closer inspection, in addition to general disorganisation medio-laterally along the folds (also as seen with WMISH) the neurons are disorganised within the apicobasal axis, several contain pycnotic nuclei (**Figures 4E,E′,E″,E‴**), and groups of dying cells appear to be aggregating under the overlying epidermis (**Figures 4E,E′,E″,E‴**, dotted circles).

Finally, we used anti-acetylated-tubulin, which labels cilia and axons, to further analyse later effects on nervous system development. In other models, inhibiting JNK signalling can cause pathfinding and arborization defects (Schellino et al., 2019). For these experiments, embryos were treated at 2.5 μM, the lowest dose of SP600125, so that they could progress past N4 to N5 and T0. In both N5 controls and treated embryos, axons can be seen throughout the medial neural canal, extending from the anterior neuropore of the cerebral vesicle all the way through the neurenteric canal connecting the neural tube to the gut (**Figures 5A^{iv},A^{v}**, red arrowhead). More lateral confocal sections further show a population of neurons on either side of the neural folds labelled with acetylated tubulin (**Figures 5A″,A‴**, red arrows). All structures are similarly labelled in T0 control and treated embryos (**Figures 5B,B′,B″**); this is particularly evident in dorsal views (**Figures 5C,C′,C″**). The most basal/ventral section taken under the floorplate and immediately above the notochord shows axons crossing the midline to the contralateral side, just posterior to the cerebral vesicle, despite cellular disorganisation and patterning defects (**Figure 5C″**). However, the branching architecture appears "messier" and axon tracts are more spread out in all treated embryos (see acetylated tubulin staining in **Figures 5A^{v},C**). We therefore cannot rule out more dramatic effects caused by more potent JNK inhibition than what we applied here.

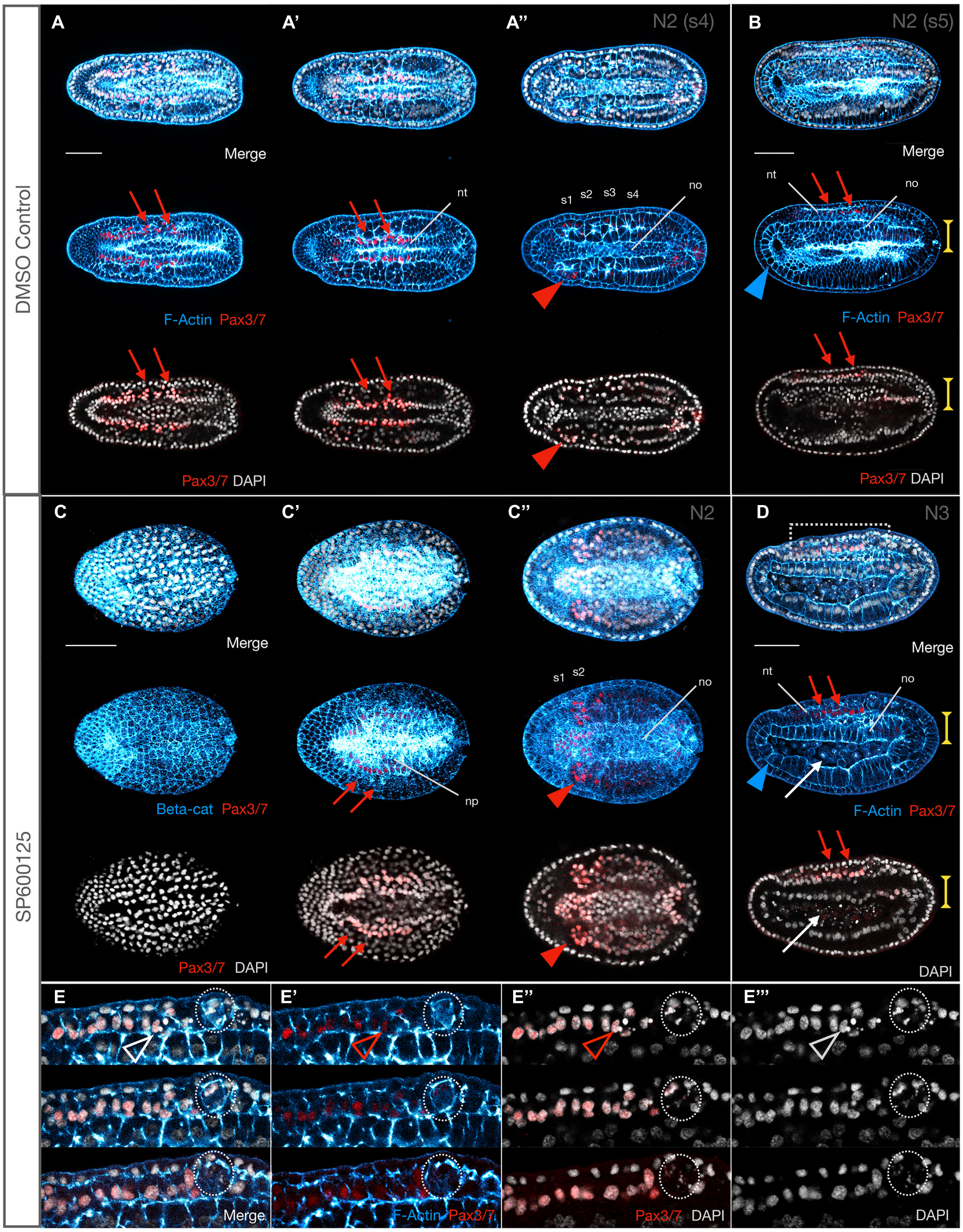


FIGURE 4 | Some Pax3/7 positive neurons are extruded and die, and somite formation is delayed after JNK inhibition. **(A,A′,A″)** Pax3/7 labels neurons along the midline (red arrows). By N2 (four somite pairs), the neural tube is effectively closed. Somites are also Pax3/7$^+$, particularly S1 (red arrowhead). All images are the same embryo in different focal planes (**A** most dorsal, **A″** most ventral). **(B)** Lateral view of a slightly more developed (five somite pairs) N2 stage embryo showing the Pax3/7$^+$ neurons of the neural tube (red arrowheads). The anterior endoderm is comprised of a monolayer of large cells (blue arrowhead). The central notochord is 3–4 cell layers thick (vertical yellow extent lines) and similarly wide **(A″)**. **(C,C′,C″)** Stage matched SP600125-treated embryos still have neural plates, with irregular Pax3/7 expression along the edges (red arrows). Only two somite pairs can be resolved, which express Pax3/7 (red arrowhead) and the notochord is 5–6 cells wide. **(D)** By the time controls reach N3, treated embryos have caught up with neural tube closure and Pax3/7$^+$ neurons are expressed along the midline (red arrowheads, compare with N2 control embryo in **B**). However, anterior endoderm is reduced (blue arrowheads) and the notochord is a monolayer of columnar cells (vertical yellow extent lines). Cells accumulating in the archenteron/forming gut are not Pax3/7$^+$ and therefore unlikely to be of neural origin. **(E,E′,E″,E‴)** Insets from embryo in **D**. Each panel consists of three confocal planes, with top being the most lateral and bottom the most medial. Some Pax3/7$^+$ neuronal cells are pycnotic, and therefore likely dying (outlined arrowheads), and several aggregate into extrusion masses (dotted outline) between the neural tube and the overlying epidermis. In all cases, anterior is to the left. Representative embryos are shown. Panels **(A,A′,A″)** and **(C,C′,C″)** are dorsal views; **(B)**, **(D)**, and **(E,E′,E″,E‴)** are lateral views. Pax3/7 immunoreactivity is in red, F-Actin or β-catenin (beta-cat) labelled membranes are in blue, and DAPI stains nuclei in grey. nt, neural tube; no, notochord; s1, s2, etc., numbered somites. Scale bars = 50 μM.

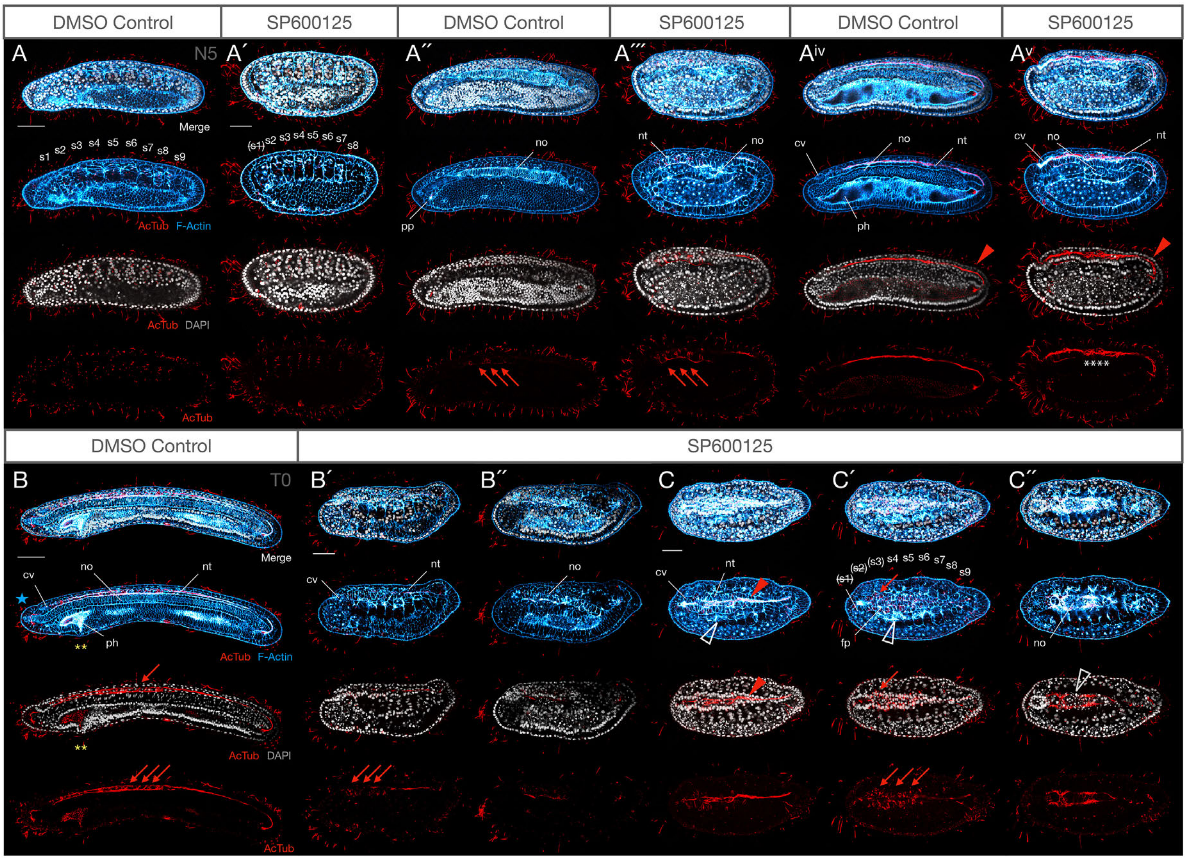


FIGURE 5 | Weak JNK inhibition allows embryogenesis to proceed to N5 and T0 stages, including axonogenesis, but mesodermal structures are still affected. **(A–A^{v})** N5 stage control and 2.5 μM SP600125-treated embryos at equivalent confocal planes. At N5, DMSO controls possess nine somite pairs **(A)**, a clear ciliated preoral pit **(A″)**, and the central canal of the nerve cord is populated by an axonal tract extending from the neuropore of the cerebral vesicle to the neurenteric canal (**A^{iv}**, red arrowhead). Treated embryos have no more than eight somites, and the first is degrading as shown by pycnotic nuclei [**A′**, indicated by (s1)]. There is no evidence of a preoral pit **(A‴)**, but peripheral axon tracts envelop neurons (**A‴**, red arrowheads) and axons extend through the neural tube **(A^{v})** similarly to controls. The notochord is highly disorganised and no pharynx proper develops. Some branching defects are evident within the central axonal tract (white asterisks). **(B–C″)** T0 stage control and 2.5 μM SP600125-treated embryos. The T0 control has more developed anterior structures, including anterior extension of the notochord past the cerebral vesicle (**B**, blue star), and cellular reorganisation of the pharynx presaging perforation (yellow asterisks). In treated embryos, the anterior endoderm remains simplified and filled with pycnotic cells **(B′,B″)** there is no extension of the notochord past the cerebral vesicle, and anterior somites continue to degenerate **(C–C″)**. Nevertheless, muscle fibres differentiate (**C,C′**, white arrowheads) and axons project contralaterally in the anterior neural tube dorsal to a pycnotic and disorganised, albeit partially differentiated, notochord (**C″**, white arrowhead). In all cases, anterior is to the left. Panels **(A–A^{v})** and **(B–B″)** are lateral views; **(C–C″)** are dorsal views. **(A,A″,A^{iv})**, **(A′,A‴,A^{v})**, **(B′,B″)**, and **(C,C′,C″)** are different confocal planes of the same embryos. Representative embryos are shown. Acetylated-tubulin immunoreactivity is in red, F-Actin labelled membranes are in blue, and DAPI stains nuclei in grey. cv, cerebral vesicle; fp, floor plate; nt, neural tube; no, notochord; s1, s2, etc., numbered somites; ph, pharynx; pp, preoral pit. Scale bars = 50 μM.

Delayed and Disorganised Somite Formation and Loss of Somite Derivatives

Expression of anterior and posterior somite markers, as well as *MRF1*, an axial muscle-specific marker, suggested that specification and differentiation proceeded at least grossly normally. However, the absence of elongation might indicate delays in somite formation, or other structural changes. F-Actin accumulation at membranes not only allows identification of somite boundaries, but it also labels the differentiated muscle fibres of the myomeres starting in late neurula stages (N4). We found that formation of the first somites by enterocoely appears to proceed broadly normally at first, although it is delayed, and the effects are concentration-dependent. For instance, by N2 stages, control embryos possess 5–6 somite pairs (median 5, $n = 7$). In contrast, 10 μM SP600125-treated embryos possess 3–4 (median 3.5, $n = 5$), revealing a significant delay in somitogenesis (Mann–Whitney $U = 0$, $P < 0.01$ two tailed). That somite specification is normal is reflected by expression of Pax3/7 in the mesoderm of the first forming somites as one would expect (Compare **Figures 4A″,C″**). Over time, somites become more and more compressed in the A-P and D-V axes, even at the lowest concentrations of 2.5 μM SP600125 (**Figures 5A,A′,B′,C′**; also **Figures 9C,D**). Interestingly, under these conditions no treated embryos allowed to progress to the N5 stage (9–10 pairs of somites, median 9.5, $n = 8$), once somite formation progresses

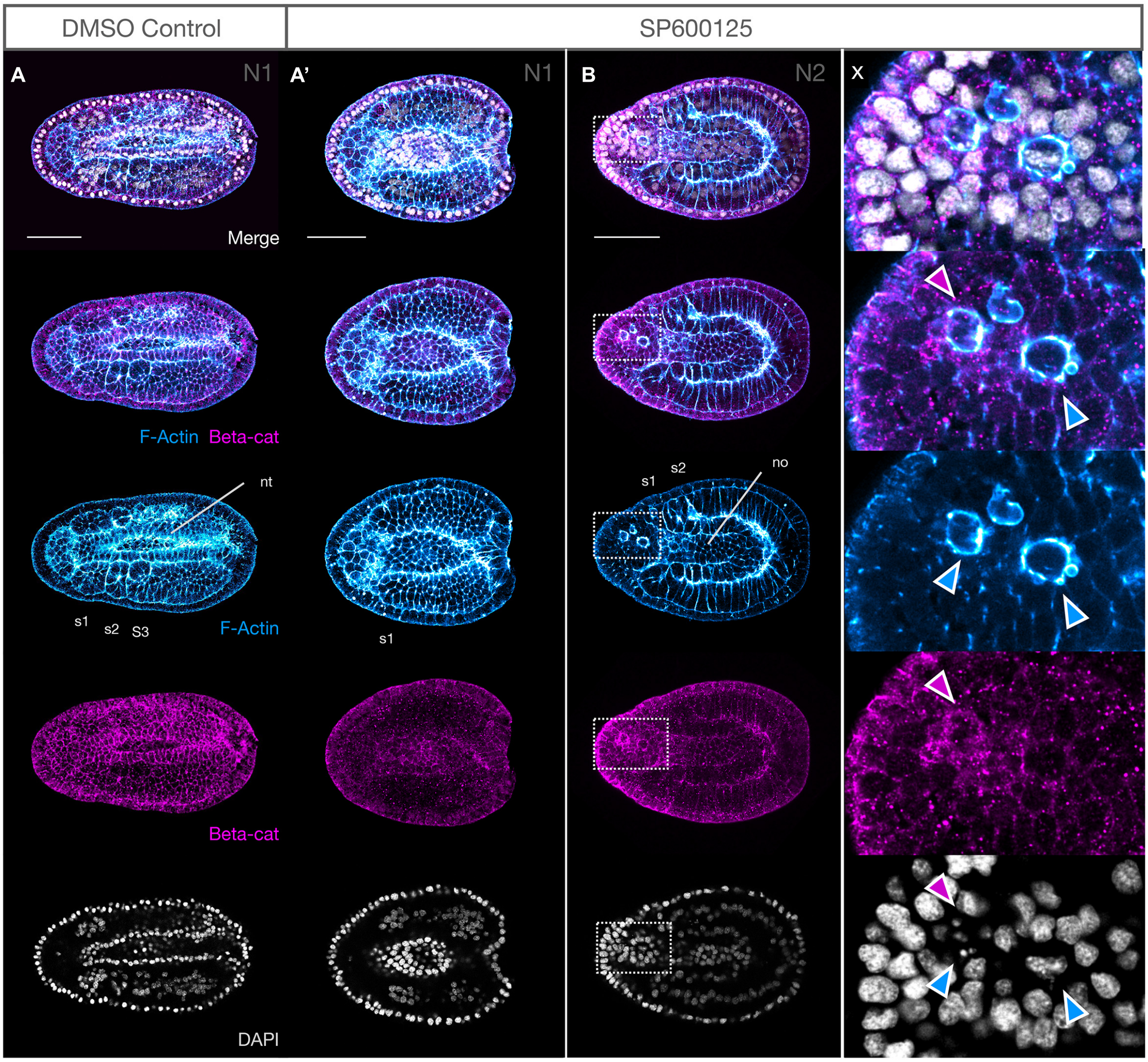


FIGURE 6 | Notochord cells round up, accumulate F-Actin and β-catenin at the membranes and undergo nuclear breakdown after JNK inhibition. **(A,A′)** At N1, the membranes of DMSO controls contain F-Actin and β-catenin, with three pairs of somites clearly outlined. The neural tube is almost entirely closed, and notochord cells can be seen anteriorly and below **(A)**. SP600125-treated embryos are delayed, with only the first somite demarcated. β-Catenin is patchily distributed but outlines notochord cells, particularly their irregular apical membranes anteriorly **(A′)**. **(B)** By N2, some notochord cells have rounded up, with strong F-Actin staining in the membranes and accumulation of β-catenin (magenta arrowheads in inset x). These cells have numerous blebs (blue arrowheads) and DAPI-stained nuclei are irregular or pycnotic (inset x). All embryos are orientated anterior to the left, with dorsal views shown. Representative embryos are shown. F-Actin and β-catenin labelled membranes are in blue and magenta, respectively, and DAPI stains nuclei in grey. nt, neural tube; no, notochord; s1, s2, etc., numbered somites. Scale bars = 50 μM.

via schizocoelic budding from the tailbud, ever possessed more than 8–9 pairs of somites (median 8, $n = 8$), even accounting for evidence that the anterior-most somite(s) might be deteriorating and dying, as evidenced by pycnotic nuclei (**Figures 5A′,C′**; also **Figures 9C,D**; Mann–Whitney $U = 0$, $P < 0.01$ two tailed). In contrast, stage matched controls possess more than 11 pairs and have passed tailbud stage T0. However, although somites are disorganised in embryos treated with SP600125, they do nevertheless appear to differentiate, as the dorsal regions contain evidence of F-Actin-rich lateral muscle fibres extending between myosepta (**Figures 5C,C′**; open white arrowheads).

Last, our assessment reveals that JNK inhibition results in the absence of formation of the preoral pit, a monociliated organ with possible homology to the vertebrate adenohypophysis (Candiani et al., 2008). During amphioxus development, it has been argued that the ventral portions of the first somite pair split off to give rise to the left and right diverticula; the left diverticulum becomes the preoral pit (or Hatschek's fossa) (Stach, 1996). Even accounting for a delay in development caused by pharmacological treatment,

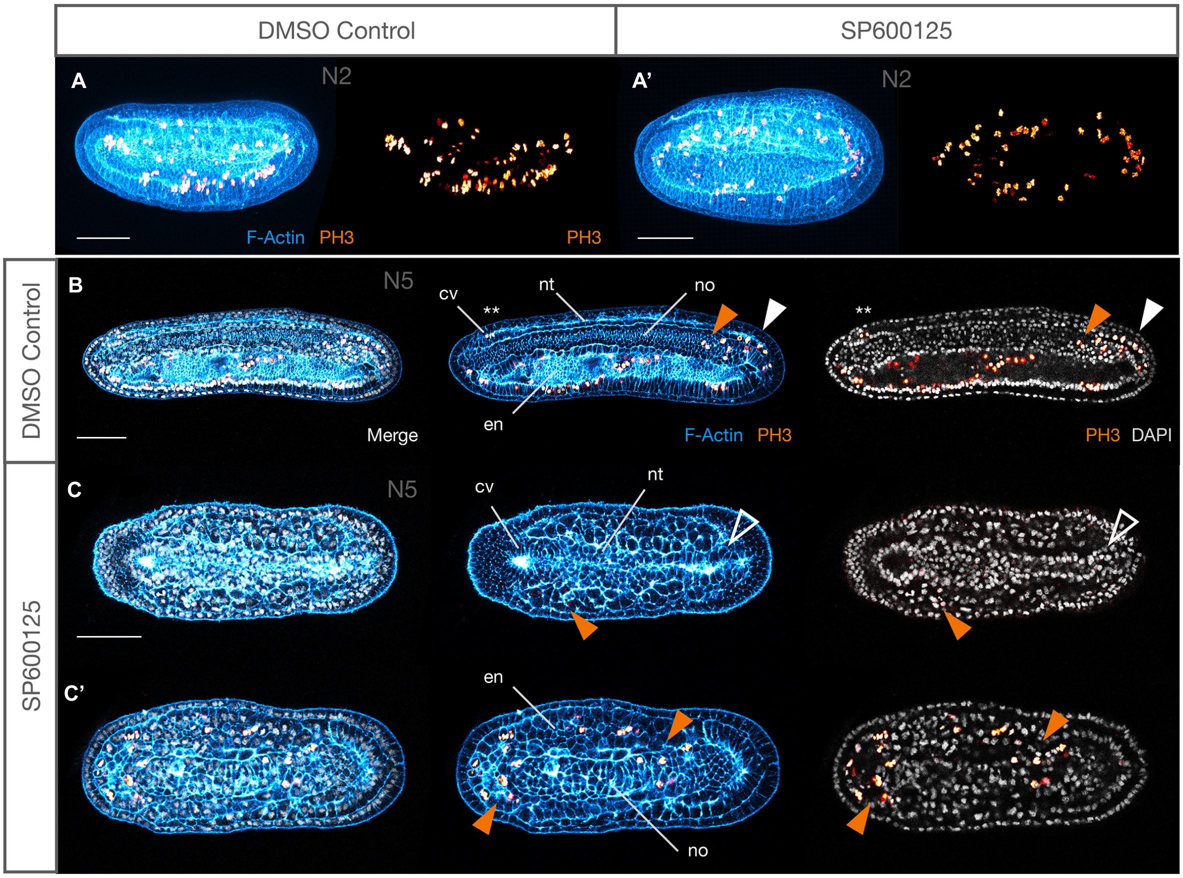


FIGURE 7 | Proliferation is maintained after JNK inhibition. **(A,A′)** Maximum projections of all confocal planes in control and SP600125-treated embryos stained for F-Actin (cyan) and phospho-histone H3 (PH3, gold) at the N2 stage showing broad distribution of mitotic cells. **(B)** Control N5 and **(C,C′)** stage matched treated embryos showing patterns of proliferation in different tissue types. In controls **(B)**, proliferation is found throughout the endoderm, but otherwise mostly localised to the posterior neural tube (white arrowhead) and notochord (orange arrowhead), and rarely in somites (not shown) or the cerebral vesicle (asterisks). In spite of dramatic size reduction in treated embryos **(C,C′)**, endoderm contains many mitotic cells as does the posterior notochord, and unusual examples of $PH3^+$ somatic cells can be seen (orange arrowheads). No neural cells divide here, even in the posterior neural tube (white outlined arrowhead). Representative embryos are shown. In all cases anterior is to the left. Panels **(A,B)** are lateral views, **(C′,C)** are dorsal views of the same embryo at different confocal planes. cv, cerebral vesicle; en, endoderm; nt, neural tube; no, notochord. Scale bars = 50 µM.

and whatever its specific origins (ventral extension of the first somite or direct pinching off from enterocoelic mesoderm), we find no evidence for the formation of a preoral pit in any of our SP600125-treated embryos (**Figure 5**). We also find no evidence of its formation and then subsequent degradation, as at no time is there a ciliated structure in this location.

Taken together, our results indicate that JNK inhibition causes a delay in the rate of enterocoelic somite formation but not of their maturation, with eventual degradation of anterior somite pairs, and an absence of preoral pit formation.

Notochord Convergence and Extension Disruption Is Accompanied by Cellular Extrusion

The observation that somite formation and differentiation is delayed led us to query the mechanism leading to the notochord elongation defect seen in SP600125-treated embryos. While shorter than N3 stage controls, the notochords of treated embryos appeared multi-layered, although some cells showed particularly strong F-Actin accumulation in membranes (**Figure 3C′**). We therefore looked at earlier stages of development to determine whether JNK inhibition caused any changes in cellular behaviour in the notochords of treated embryos. Indeed, along with a delay in somitogenesis, notochords of treated embryos are both broader (**Figure 4C′′**) and dorso-ventrally shallower (**Figure 4D**, yellow extent lines) than their N2 controls (compare with A′′ and B). In other words, the notochord fails to extend in the antero-posterior and dorso-ventral axes in a classic convergence-extension phenotype, remaining as a monolayer even in N3 stage embryos. Careful analysis of embryos at N1 (**Figure 6A,A′**) and N2 stages (**Figure 6B** and inset x) indicates that some cells begin to show irregular contours at N1, followed by rounding up and blebbing. Membranes accumulate F-Actin as well as β-catenin (**Figure 6B** and inset x), and nuclei begin to fragment (inset x, note DAPI staining in rounding cells).

Even with weaker treatments with SP600125 (2.5 μM) and allowing embryos to develop to N5 or T0 stages, the notochords remain disorganised relative to those of controls, lacking the classic "stacked coin" morphology of their matched controls (**Figures 5A^{iv},B** vs. **A^{v},B$''$**). By these late stages, many notochord cells are pycnotic (**Figure 5C$''$**, grey outlined arrowhead) and there is no anterior notochord extension past the cerebral vesicle as seen in T0 controls (**Figure 5B**, blue star).

Taken together our results suggests that a delay in differentiation is not sufficient to account for the CE defects seen in embryos after JNK inhibition. Cellular extrusion starting early on also accounts for the shorter, disorganised notochords seen after SP600125 treatment.

Changes in Proliferation Are Not the Prime Mediator of Size

JNK is known to play an important role in cellular survival, as well as regulation of the cell cycle (Pinal et al., 2019; La Marca and Richardson, 2020). We hypothesised that growth and patterning defects might be caused by dysregulation of the balance between cell death and proliferation, either through a broad reduction in proliferation or through changes in patterns of cell division in different tissues. We therefore compared proliferation between treated and control embryos using an antibody against PH3, a marker commonly used to detect late G2 and mitosis phases of the cell cycle. Since embryos at higher levels of drug concentration (10 μM), or those allowed to develop for a long period even at lower concentrations (2.5 μM) showed extreme changes in morphology and reduction in size, complicating interpretation, we first assessed proliferation in N3 stage control and 5 μM – treated embryos, an interval permitting sufficient time for the drug to act without causing major cellular shedding into the archenteron/gut cavity.

Although total cell number is difficult to accurately quantify in entire amphioxus embryos, maximum projection of all sections containing PH3-positive cells shows a broadly similar distribution and density of proliferating cells (representative embryos in **Figures 7A,A$'$**). By N5, there is a clear difference in size of the treated embryos, which consist of absolutely many fewer cells. Nevertheless, comparison of proliferation patterns again does not indicate any clear differences. In the control, PH3^{+} cells are found in the cerebral vesicle and somites, though rarely, as well as throughout the endoderm, and are particularly prevalent in the tailbud region, including posterior nerve cord and notochord (**Figure 7B**). There is little to no proliferation in the rest of the nerve cord or notochord. This, and the inter-individual variability in proliferation patterns, is consistent with what has been observed in other studies (Holland and Holland, 2006; Carvalho et al., 2021). Similarly, mitotic indices are rare in the nerve cord, along the notochord axis and in the cerebral vesicle in embryos treated with SP600125. However, considerable proliferation can be seen in both anterior and lateral endoderm, as well as in the posterior of the notochord (**Figures 7C,C$'$**). Taken together, our results suggest that proliferation does continue throughout the tissues, even after a long exposure to the JNK inhibitor, but may not be sufficient to compensate for cellular losses. However, deeper analysis of cell cycle dynamics in large sample sizes will be required to confirm these initial observations.

Inhibiting JNK Signalling Causes Caspase-3-Dependent Apoptosis

The observation that the large number of cells within the archenteron or between tissues often contained pycnotic nuclei suggested that JNK-mediated loss of polarity might be associated with the activation of cell death pathways. Caspases are typically activated during apoptosis and can be detected using an antibody against cleaved Caspase-3. We found a strong and persistent expression of anti-cCasp-3 in all embryos treated with the JNK inhibitor SP600125 at all concentrations used and at every stage tested (**Figure 8B**). In contrast, immunoreactivity was never detected in tissues of control animals (**Figure 8A** and **Supplementary Figure 4**), with the exception of an occasional shed cell within the gut lumen in N4-T0 stages (e.g., **Figures 9A$'$,A$''$**). Unexpectedly, even in treated embryos, cells within tissues very rarely expressed the apoptosis marker and it was mostly confined to cells within the extracellular spaces within the archenteron/gut, in the extruded "dorsal fin" cells, or within the open space between the nerve cord and overlying epidermis (**Figure 8B**, magenta arrows and arrowhead). In early neurula stages, cells from the gut lumen may have squeezed through the neurenteric canal while still open. However, examples of isolated cases of cCasp-3^{+} pycnotic nuclei in CNS and chordamesoderm cells undergoing extrusion in strong early treated embryos suggest that there may be several origins of delaminated cells. We never saw an example of cCasp-3 expression in gut cells proper, but rounding of anterior ventral endoderm cells, particularly in strong early treated embryos, and accumulation of such cells within the anterior gut suggest that JNK inhibition might also lead to delamination of these cells and their subsequent death (**Figure 8B**, blue arrows and arrowhead). Taken together, these data suggest that Caspase-3-mediated apoptosis may at least in part be a secondary effect of loss of polarity and positional identity.

Pharmacological Inhibition of Caspase Activity Partially Rescues the *orca* Phenotype

Loss of JNK activity might either directly activate Caspase-dependent apoptosis, or alternatively, indirectly cause apoptosis in cells that lose polarity and leave their assigned tissue. In order to attempt to distinguish between the two hypotheses, we performed a series of experiments to determine whether or not we could "rescue" the *orca* phenotype by co-treating embryos with SP600125 and the pan-Caspase inhibitor Z-VAD-fmk and comparing with individual treatments. Z-VAD-fmk is a pan-Caspase inhibitor that has been shown to block amphioxus Caspase-3/7 activity *in vitro* (Bayascas et al., 2002), but the phenotypes were never assessed *in vivo*. Because of the strength of the phenotype and the need to increase the developmental time window to assess phenotypic effects, we combined a strong Z-VAD-fmk treatment (200 μM) with

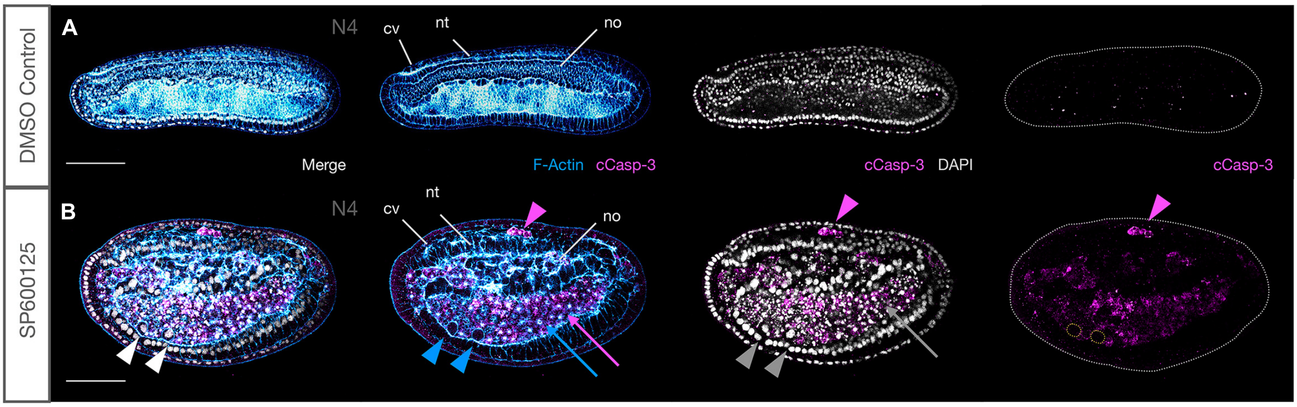

DMSO Control
SP600125
A
N4
cv
nt
no
Merge
F-Actin cCasp-3
cCasp-3 DAPI
cCasp-3
B
N4
cv
nt
no

FIGURE 8 | JNK inhibition induces Caspase-3-mediated apoptosis. **(A)** Under ordinary control conditions, there is no Caspase-3-mediated apoptosis in N4 embryos. **(B)** JNK inhibition with SP600125 causes dramatic accumulation of cleaved Caspase-3 in cells that have been extruded between the neural tube and epidermis (magenta arrowhead), around the notochord and within the accumulated cells of the archenteron where pycnotic nuclei (grey arrow) and F-Actin aggregates (blue arrow) are observed. Rounded cells that have delaminated with strong membrane staining and no F-Actin aggregates have intact nuclei and are cCasp-3⁻. Note how ventral cells highlighted by white, blue, and grey arrowheads lack cCasp-3 staining in single magenta channel (cells outlined with yellow dotted lines). Representative embryos are shown. Phalloidin staining of F-Actin is in blue and cleaved Caspase-3 (cCasp-3) immunoreactivity is in magenta, while DAPI stained nuclei are in grey. Embryos are orientated with anterior to the left and dorsal up. cv, cerebral vesicle; nt, neural tube; no, notochord. Scale bars = 50 μM.

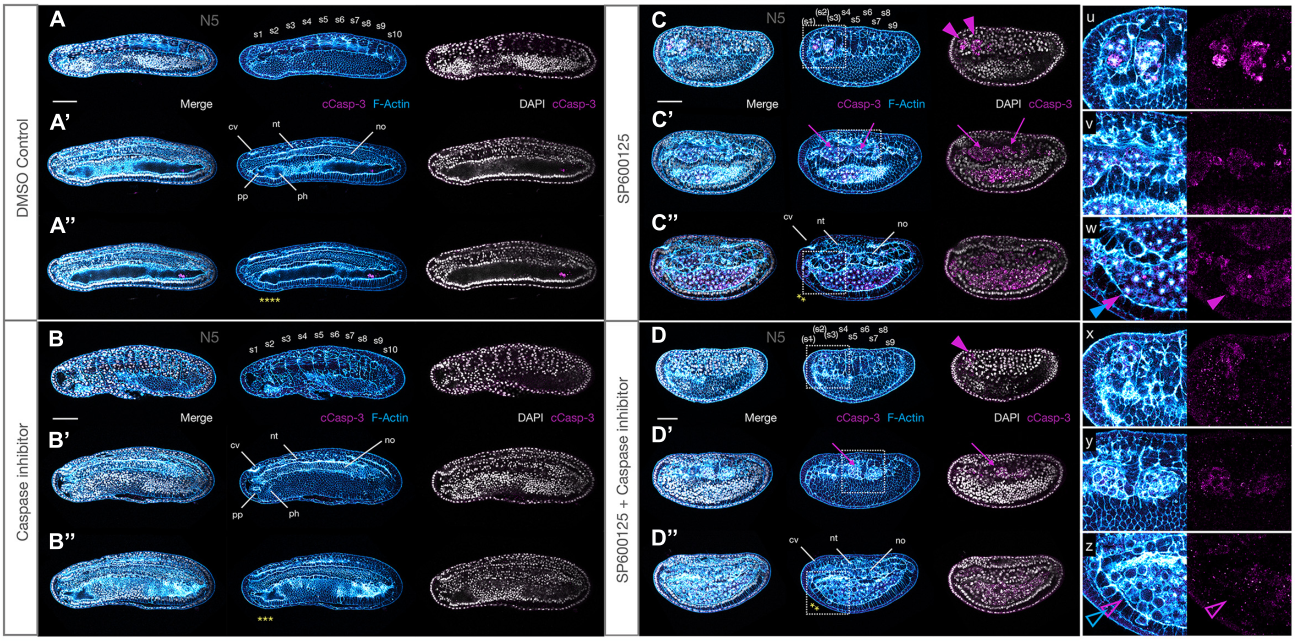

DMSO Control
Caspase inhibitor
SP600125
SP600125 + Caspase inhibitor
A
N5
s1 s2 s3 s4 s5 s6 s7 s8 s9 s10
Merge
cCasp-3 F-Actin
DAPI cCasp-3
A'
cv
nt
no
pp
ph
A''
****
B
N5
s1 s2 s3 s4 s5 s6 s7 s8 s9 s10
Merge
cCasp-3 F-Actin
DAPI cCasp-3
B'
cv
nt
no
pp
ph
B''
***
C
N5
(s1) (s2) (s3) s4 s5 s6 s7 s8 s9
Merge
cCasp-3 F-Actin
DAPI cCasp-3
C'
C''
cv
nt
no
**
D
N5
(s1) (s2) (s3) s4 s5 s6 s7 s8 s9
Merge
cCasp-3 F-Actin
DAPI cCasp-3
D'
D''
cv
nt
no
**
u
v
w
x
y
z

FIGURE 9 | A pan-Caspase inhibitor only partially rescues weak JNK inhibition phenotypes. **(A,A′,A″)** N5 stage DMSO control with 10 somite pairs **(A)**, a preoral pit **(A′)**, and expanding pharyngeal region (**A″**, yellow asterisks). **(B,B′,B″)** Stage matched embryo treated with the pan-Caspase inhibitor Z-VAD-fmk at high concentration (200 μM). Embryos are slightly shorter but otherwise equivalent to the controls, with all differentiated structures present. **(C,C′,C″)** Weak JNK inhibition with 2.5 μM SP600125 shows dramatic accumulation of cleaved Caspase-3 in dying cells, including degenerating anterior somites (**C**, s1–s3, magenta arrowheads; inset u), cells extruded around the notochord (**C′**, magenta arrows, inset v), and within the archenteron (**C′**, inset w). No real pharyngeal region has differentiated (yellow asterisks). **(D,D′,D″)** Embryos co-treated with SP600125 and Z-VAD-fmk are grossly similar to embryos treated with SP600125 alone. However, cleaved Caspase-3 immunoreactivity is significantly reduced both quantitatively and qualitatively. The notochord, although still delayed, has a more regular cellular arrangement (compare **C″,D″**), and cellular protrusions (insets x and y) contain rounded cells, but fewer central F-Actin aggregates. Rounded cells in the anterior of the archenteron have a similar "healthier" phenotype (compare w and z, filled and outlined blue and magenta arrowheads), with evidence that at least some may originate from within the anterior ventral endoderm layer. Phalloidin staining of F-Actin is in blue and cleaved Caspase-3 (cCasp-3) immunoreactivity is in magenta, while DAPI stained nuclei are in grey. Embryos are orientated with anterior to the left and dorsal up. Each panel series shows one representative embryo at different confocal planes. cv, cerebral vesicle; nt, neural tube; no, notochord; s1, s2, etc., somite pairs. Scale bars = 50 μM.

the lowest 2.5 μM SP6000125 treatment condition, allowing embryos to develop with clearly discernible morphology to N4 or even T0 stages.

We found no overt anatomical differences between Z-VAD-fmk and control embryos (compare **Figures 9A,A′,A″** with **B,B′,B″**), except perhaps a negligible decrease in size under treatment conditions. Similarly, on first inspection, JNK inhibition alone vs. SP600125 + Z-VAD-fmk inhibitor appeared to generate similar phenotypes (compare **Figures 9C,C′,C″** with **D,D′,D″**). However, both the nerve cord and notochord appeared more regular and extended, and cCasp-3 expression was reduced in the double-treated embryos, both quantitatively (fewer patches) and qualitatively (lower intensity) (see insets u–w and x–z), suggesting that Z-VAD-fmk might provide some mild rescue of the Caspase-dependent apoptosis induced by SP600125. In addition, although both conditions resulted in rounded cells accumulating within the gut space, as well as lateral bulges of rounded actin-rich cells between the somites and endoderm (insets v and y), many fewer of these were pycnotic, with the characteristic central actin aggregates, particularly within the anterior gut when both inhibitors were used (compare closed and open magenta and blue arrowheads in w and z, respectively). Indeed, these cells appear to have rounded up and exited their tissues, but not to have died. Taken together, the mild rescue seen in these experiments suggest that the *orca* phenotype is probably the result of both the effect of JNK inhibition on polarity/cell movements and on the normal programmed cell death required to remodel amphioxus tissues during early development.

MAPK/ERK Signalling Is Ectopically Activated

Having ascertained that JNK inhibition leads to an apparent delamination and death of misplaced cells in several tissues, resulting in CE defects, we wanted to determine the mechanism by which this might occur. Several reports suggest that JNK interacts with several pathways including MAPK and TGFβ/BMP to coordinate cellular movements, cell division and polarity (Gros et al., 2010). Specifically, regulation of MAPK–ERK activity governs cellular behaviour and extrusion, followed in many cases by apoptosis, in vertebrate and invertebrate models (Moreno et al., 2019; Aikin et al., 2020; Tada, 2021). SP600125 is a highly selective inhibitor of c-Jun phosphorylation but has no observable effect on ERK or p38 phosphorylation (Bennett et al., 2001). Thus, any changes in ERK phosphorylation act as a readout of MAPK signalling, and we therefore reasoned that this pathway might be responsible for some of the altered cellular behaviour and patterning that we observed. We therefore assessed levels of pERK in SP600125-treated N2 stage embryos, prior to the onset of the most dramatic defects caused by JNK inhibition, using a pERK antibody that detects active ERK signalling. Our results indicate that under control conditions, there was very little localised expression of activated MAPK (**Figures 10A,A′**). However, in embryos treated with the JNK inhibitor, we found strong ectopic expression in epidermal cells, and in a few cases, in neural cells (**Figure 10B′** and inset y). More dramatically, we saw an unexpected upregulation of pERK in many medial notochord cells, but also in anterior notochord cells (**Figure 10B** and inset x, white outlined rainbow arrowheads), specifically in cells neighbouring the rounding cells, whose membranes are F-Actin rich, and which appear to be targeted for extrusion (**Figure 10B** and inset x, blue outlined black arrowheads). This response was also apparent at later stages (N3) when most cell rearrangement occurs within the notochord (note multi-layered aspect) and apoptotic cells have accumulated within the gut (**Figure 10C** and inset z). Taken together, these data suggest that there is a correlation between loss of JNK-mediated polarity, cellular extrusion and pERK activation, and that this response is non-cell autonomous in the chordamesoderm/notochord.

DISCUSSION

Although pharmacological inhibition of JNK activity using SP600125 is commonplace in vertebrate cell lines or epithelia *in vitro* (Yu et al., 2019; Brunt et al., 2021), studies on whole embryos are comparatively rare. This is surprising given its versatility as a tool, particularly in aquatic organisms with external development, where it allows studies of morphogenetic processes of complex tissues within specific temporal windows, without the need for designing labour and time-intensive tools. This is true even in models with long-established genetic manipulation methodologies such as zebrafish (e.g., He et al., 2016; Li et al., 2019), or even ascidians (e.g., Krasovec et al., 2019), especially where late functions need to be assessed. In amphioxus, treatment with SP600125 results in a complex phenotype likely encompassing the full range of its functions, suggesting it is a good chordate model for understanding the evolution of JNK function at the invertebrate to vertebrate transition.

Evolution of the Chordate Functional JNK Repertoire

The two whole genome duplications that occurred in the vertebrate ancestor (Dehal and Boore, 2005) resulted in the presence of multiple paralogues in vertebrate relative to invertebrate chordate genomes, barring any lineage-specific duplication or gene loss events. The evolution of JNKs represents a near textbook-case scenario, with single *jnk* genes in amphioxus and ascidians (similarly to other invertebrates, e.g., *basket* in *Drosophila*; Satou et al., 2003; Bertrand et al., 2009), and three orthologues in mammals (Zeke et al., 2016). In parallel, the mitogen activated signalling cascade upstream of JNK is considerably more simplified in invertebrate chordates relative to vertebrates (Satou et al., 2003; Bertrand et al., 2009). In vertebrates, such duplication events have conferred some genetic redundancy, protecting critical cellular functions. For instance, single *JNK* mutant mice survive, but double mutants for *JNK1* and *JNK2* are embryonic lethal, with region-specific dysregulation of apoptosis in the early brain (Kuan et al., 1999). Apparently, some constraint exists on the number of JNK copies tolerated in vertebrate genomes. Bony fish, for instance, might be expected to harbour many more duplicates (at least 8, assuming

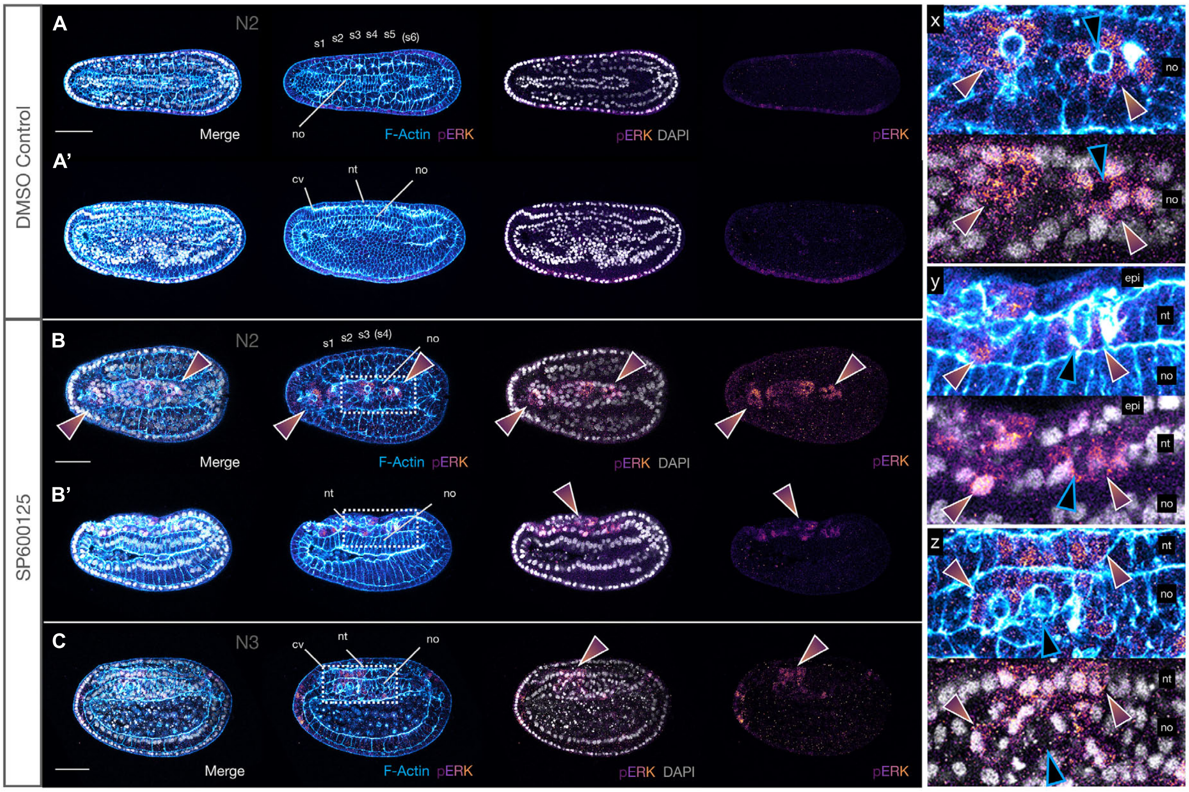


FIGURE 10 | JNK inhibition leads to upregulation of pERK in neighbours of extruding cells. **(A,A′)** N2 stage DMSO control embryo with five formed somite pairs. The notochord cells are arranged in two parallel rows, except anteriorly under the cerebral vesicle, and are multiple cell layers thick. There is little to no pERK staining. **(B,B′)** SP600125-treated embryos allowed to progress to the same stage have broader monolayered notochords. They accumulate cytoplasmic and nuclear pERK in notochord (rainbow arrowheads in **B,B′** and inset y) and overlying neural and epidermal cells (rainbow arrowheads in **B** and inset x). Two different embryos are shown. **(C)** This pattern is seen in treated embryos allowed to progress to the equivalent of the N3 stage, in which notochord cells appear to attempt to stack dorsoventrally in multiple layers (inset z). In all cases, pERK accumulation occurs in cells neighbouring rounding cells apparently in the process of extruding (note apical rounding and F-Actin accumulation in membranes, blue outlined black arrowheads in insets x, y, and z). Phalloidin staining of F-Actin is in blue and phosphorylated ERK (pERK) immunoreactivity is in a purple/orange gradient, while DAPI stained nuclei are in grey. Embryos are orientated with anterior to the left either dorsal (**A,B** and inset x) or lateral views (**A′,B′,C** and insets y and z) are shown. Representative embryos are shown. cv, cerebral vesicle; epi, epidermis; nt, neural tube; no, notochord; pharynx, ph; pp, preoral pit; s1, s2, etc., somite pairs. Scale bars = 50 μM.

no gene losses) due to an independent genome duplication in the teleost ancestor (Glasauer and Neuhauss, 2014). Surprisingly, however, a single *JNK2* and *JNK3*, and only two *JNK1* paralogues named *JNK1a* and *JNK1b* have been identified in zebrafish (Santos-Ledo et al., 2020). Similarly, *Xenbase* reports only two *MAPK8/JNK1*, one *MAPK9/JNK2*, and two *MAPK10/JNK3* genes in *Xenopus laevis* (Karimi et al., 2018). This is still fewer than expected but can be explained by the species' allotetraploid origins, likely reflecting asymmetric evolution of the subgenomes, with one chromosome set more prone to gene loss (Session et al., 2016). To compensate, alternative splicing of *jnk*s appears to be rampant in vertebrates, increasingly dramatically the functional repertoire of JNK proteins. For example, in zebrafish, four isoforms have been identified for each of the JNK1a and JNK1b paralogues, with the *jnk1a Ex7 Lg* transcript required to generate correct numbers of cardiac progenitors within the ventricle of the developing heart (Santos-Ledo et al., 2020). In humans, all three JNK genes encode multiple alternative splice isoforms: four each for JNK1 and JNK2, and eight for JNK3, of which only three have been characterised (Zeke et al., 2016). The isoforms include alternative exons in the C-terminal lobe of the kinase domain (α- and β-isoforms arising from use of a mutually exclusive exon pair, normally the sixth, but nomenclature is inconsistent), and in both the N- and C-terminal flexible extensions of the proteins. Functionally, it is predicted that different JNK isoforms will exhibit different kinetics and enzyme activities as well as substrate affinities (Zeke et al., 2016). Intriguingly, Ait-Hamlat et al. (2020) have recently predicted the existence of a new JNK1 isoform lacking exons 6, 7, and 8, the critical region for kinase structural stability and MKP7 phosphatase binding, which they named JNK1δ. Nevertheless, the isoform has stable secondary structure in solution, is expressed in human, mouse and fugu, and produces four peptides (Ait-Hamlat et al., 2020). Mechanistically, splicing of JNKs is regulated by Nova through differential availability of Nova-binding sites in the paralogues (Letunic et al., 2002; Zeke et al., 2016), and Nova also regulates JNK splicing in amphioxus in a tissue-specific fashion (Irimia et al., 2011). These studies highlight the increased functional complexity possible through both duplication and alternative splicing, and help explain why it is difficult to dissect the myriad roles of JNK signalling even in amphioxus, which has a single gene. It will be informative to further study differences between amphioxus and vertebrates in function of JNK paralogues vs. isoforms during development through more targeted approaches.

JNK Plays a Conserved Role in Gastrulation and Convergence Extension Movements

This study adds to several lines of evidence pointing to a conserved role for JNK signalling in regulating tissue movements, both during gastrulation and in processes intimately linked to axis elongation. For instance, our gastrulation arrest phenotype, with little to no change in cell type specification, closely parallels that seen in sea urchins treated with SP600125, which inhibited invagination, but not endoderm differentiation (Long et al., 2015). Although our understanding of early patterning processes in amphioxus has advanced dramatically in the last few years (Li et al., 2017; Aldea et al., 2019; Kozmikova and Kozmik, 2020; and many others), still little is known about the invagination process at the cellular level. While we focus here on the role of JNK in later processes, it is intriguing that we found no evidence of inductive events between endoderm and ectoderm, since disruption of archenteron formation and lack of normal apposition of endoderm and ectoderm resulted in no changes in gene expression; notochord, mesoderm and neural tissues all appear properly specified in early treatments, at least until late gastrula stages.

Intimately linked to axis elongation in chordates is the process of notochord morphogenesis, which like invagination requires precise spatio-temporal control of cellular behaviours. In vertebrates such as zebrafish and *Xenopus*, correct CE movements of dorsal mesoderm at the midline require JNK signalling downstream of Wnt/PCP pathways (Yamanaka et al., 2002; Kim and Han, 2005; Williams and Solnica-Krezel, 2020), although there is considerable variation in how mediolateral intercalation behaviour is regulated across species (Keller and Sutherland, 2020). Nevertheless, our results suggest conservation of the core polarity mechanism in chordates, since inhibition of JNK activity causes dramatic defects in medial chordamesoderm and notochord cells, in spite of apparent correct specification of midline structures (e.g., *Brachyury2* and *chordin*), and a strong axial elongation phenotype. Intriguingly, this also translates to a truncation of the anterior notochord, correlating with a reduction in *six3/6* expression in a notochordal domain rostral to the hypothalamo-prethalamic primordium of the cerebral vesicle ("HyPTh"; Albuixech-Crespo et al., 2017). Despite several differences with vertebrate anatomy and genoarchitecture, this *six3/6*$^{+}$ region may represent the amphioxus homologue of the prechordal plate (Albuixech-Crespo et al., 2017). In vertebrates, notochord formation from axial mesoderm requires the complex orchestration among active migration of prechordal cells (Bosze et al., 2020), non-canonical Wnt11-mediated cellular movements of paraxial mesoderm (Heisenberg et al., 2000), and CE of deep neural cells (Keller and Sutherland, 2020). Treatment with SP600125 also inhibits the mediolateral cell intercalation of the neurectoderm required for axial elongation in the annelid *Platynereis dumerilii* (Steinmetz et al., 2007). Thus, inhibition of JNK throughout amphioxus tissues likely perturbs cellular polarity and coordination of behaviours across germ layers, culminating in the truncated axis phenotype we see here, highlighting an ancient global function for this pathway in regulating CE movements.

JNK Is Required for Formation of Oral Structures in Amphioxus

Our results show that inhibiting JNK signalling has a profound and specific effect on the formation of anterior and oral structures in amphioxus, beyond those attributed to anterior notochord extension defects. For instance, we show that even mild treatments with SP600125 result in loss of the ciliated preoral pit, derivative of the initial kidney (or Hatschek's left diverticulum) and by extension cells of the posterior first left somite (Holland, 2018). This may be due to a loss of migratory activity of precursors detaching from the somite, and/or to degradation of the somite itself. We also fail to see any of the oral perforations in pharyngeal endoderm that presage club shaped gland, gill slit or mouth formation. In fact, cells appear to continue to round up and delaminate from anterior ventral endoderm until very late developmental stages (e.g., **Supplementary Figure 4**). Although there is continued debate surrounding the homology of chordate mouths (reviewed in Holland, 2018), the conservation of a role for JNK in perforation at the junction of epithelial membranes is consistent with our observations. In *Xenopus*, JNK signalling is crucial for epidermal integrity and E-Cadherin localisation at the adherens junction, with a failure of buccopharyngeal opening and subsequent embryonic mouth formation in embryos treated with SP600125 or injected with morpholinos against *JNK1* (Houssin et al., 2017). However, this phenotype was not caused by either the apoptotic or cell cycle regulatory functions of JNK signalling (Houssin et al., 2017). Because our treatments were continuous rather than targeted to the specific window of pharyngeal differentiation, we cannot exclude the possibility that other factors are at play. Indeed, even embryos treated with only 2.5 μM SP600125, with or without Caspase inhibition, continue to show large numbers of rounded cells apparently delaminating within the pharyngeal region. This is also consistent with a more general function in adhesion or anterior endoderm specification, similarly to the role played by Wnt/JNK signalling in determining foregut identity and cellular morphology in zebrafish (Zhang et al., 2016) and endodermal tissue integrity during gut elongation in *Xenopus* (Dush and Nascone-Yoder, 2013). The apparent loss of the anterior ventral endoderm expression domain of *wnt11* might be consistent with a loss of specific structures in this region. Future work should be aimed at specifically addressing the link between JNK signalling and morphogenesis of oral and "head" structures, including the expression of late endoderm markers, in cephalochordates.

Wnt Genes May Be Transcriptional Targets of JNK Signalling in Amphioxus

In several cellular contexts, the non-canonical Wnt/PCP pathways are mediated by JNK *via* specific Wnt ligands, independently of nuclear β-catenin-dependent canonical Wnt signalling. As previously discussed, Wnt/JNK signalling through Wnt5 or Wnt11 paralogues plays an important role during avian, zebrafish and *Xenopus* gastrulation (Heisenberg et al., 2000; Hardy et al., 2008; Seo et al., 2010; Keller and Sutherland, 2020). Wnt5a is also responsible for correct orientation of mesenchymal cell movements and division during proximo-distal limb bud elongation in chick and mouse through JNK signalling (Gros et al., 2010). However, Wnt genes may also themselves be targets of JNK activity. For instance, depletion of JNK activators MKK4b and MKK7 results in CE defects during gastrulation and abnormal somitogenesis in zebrafish, with *wnt11* a direct downstream target of Wnt/JNK signalling (Seo et al., 2010).

In amphioxus, the functional data permitting the identification of non-canonical Wnt ligands are still lacking, and comparative studies of the Wnt family expression atlas across chordates suggest that considerable function shuffling of Wnt ligands has occurred in these lineages (Somorjai et al., 2018). In particular, the expression of *wnt2* in the amphioxus cerebral vesicle/nerve cord is conserved with vertebrates but is a cephalochordate novelty in the notochord (Somorjai et al., 2018). Nevertheless, our data suggest that *wnt2* and *wnt11* may be (direct or indirect) targets of a non-canonical Wnt/JNK signalling pathway in amphioxus within specific spatio-temporal contexts. For instance, *wnt11* appears to be ectopically activated in the posterior ectoderm after SP600125 treatment, which is consistent with the zebrafish data (Seo et al., 2010). However, it is also downregulated in the paraxial mesoderm and in the chordoneural hinge, and possibly in the ventral endoderm. This could be explained if Wnt11 is both upstream of JNK signalling and its target, with SP600125 resulting in threshold-dependent autoregulation in a negative feedback loop. This is not unheard of, as we have shown similar contradictory downregulation of *sp5* in the amphioxus tailbud in response to strong and prolonged treatment with a GSK3-β inhibitor (which activates Wnt/β-catenin signalling), when otherwise it behaves as a positive target (Dailey et al., 2017). *Wnt2*, in contrast, appears to be downregulated persistently and exclusively in the notochord, with no effect of the inhibitor on expression in the nerve cord, suggesting *wnt2* may be a direct target of JNK signalling in chordamesoderm. However, we cannot currently exclude a second possibility, which is that the loss of *wnt2* and *wnt11* expression reflects an indirect effect of JNK on notochord or endoderm differentiation, respectively. In vertebrates, Wnt2, Wnt5a/5b, and Wnt11 induce differentiation of human cardiomyocytes from lateral plate mesoderm *via* an atypical Wnt/Ca^{2+} pathway (Mazzotta et al., 2016). There is also increasing evidence for direct crosstalk between JNK and the canonical Wnt signalling pathways at the transcriptional level (Bikkavilli and Malbon, 2009), complicating interpretation. Because of its late onset of expression, $Wnt2^{-/+}$ or $Wnt2^{-/-}$ transgenics could conceivably be generated in amphioxus to help distinguish between these possibilities.

Changes in Caspase-3-Mediated Apoptosis and Proliferation Within Tissues Cannot Explain SP600125-Induced Cellular Loss

It is now well established (if not well understood) that JNK plays important pro-proliferative and pro-apoptotic roles during development and tissue replacement, depending on the cellular context (reviewed in Ricci and Srivastava, 2018; Pinal et al., 2019; Guerin et al., 2021). JNK activity has also been proposed to act as a pivot point, balancing cell death and proliferation to maintain body proportions during whole body regeneration and starvation-induced homeostatic degrowth in flatworms (Almuedo-Castillo et al., 2014). Our *orca* phenotype could therefore be interpreted as a dysregulation of this equilibrium, resulting from either constitutive or local increases in apoptosis or, alternatively, decreases in proliferation.

In amphioxus, TUNEL staining – an assay used to detect apoptotic DNA cleavage as well as other types of DNA damage – previously suggested that programmed cell death is regulated during embryogenesis, with apoptosis likely required for anterior endoderm remodelling starting during neurulation stages (Bayascas et al., 2002). At larval stages, dying cells are scattered throughout the tissues, with concentration of signal in the pharyngeal and tail regions (Bayascas et al., 2002), and a burst is again seen at the onset of metamorphosis around the forming mouth, gills and anus, in the club shaped gland and preoral pit, and in tail epidermis (Holland et al., 2009). Tail regression, on the other hand, appears to primarily employ remodelling (Koop et al., 2011). In contrast, in the solitary ascidian *Ciona*, programmed cell death is necessary for regression of the tadpole tail and settlement (Krasovec et al., 2019). Here, signalling by both ERK and JNK MAP kinases is required (Chambon et al., 2007), and tail regression fails when either the pathway is inhibited, or Caspase activity is blocked (Chambon et al., 2002; Krasovec et al., 2019).

Amphioxus is known to have a single Caspase-3/7 orthologue of mammalian effector caspases with similar substrate preference to vertebrate Caspase-7, and which induces apoptosis. Its transcript expression is rather non-specific within mesendoderm and uncorrelated with TUNEL; the authors conclude that the role of Caspase-3/7 is to confer the ability for cells to die rather than instructing apoptosis (Bayascas et al., 2002). Similarly, in control embryos we see little convincing immunoreactivity with cCasp-3 antibody, which marks cells that are committed to apoptosis. In contrast, SP600125 induces considerable nuclear fragmentation coincident with expression of cCasp-3. However, $cCasp3^{+}$ cells are rarely found within the embryonic tissues, being generally limited to cells that have already clearly delaminated into luminal or coelomic spaces (e.g., within the archenteron or between the epidermis and neural tube), and only within specific early developmental windows corresponding to neurula stages. Therefore, we do not see a pattern of Caspase-3 activity corresponding to previously reported TUNEL staining. This suggests that other cell death pathways (possibly including necroptotic; Tait and Green, 2008) may be involved during normal development. Consistent with this, single treatments with the pan-Caspase inhibitor Z-VAD-fmk, which was demonstrated to inhibit Caspase-3/7 from cleaving its substrates *in vitro* (Bayascas et al., 2002), had little effect on morphogenesis. However, it did appear to rescue to some degree the Caspase-3-mediated apoptosis of cells that had already delaminated into luminal spaces. Although we have no lineage data to support this (indeed, we never saw gene marker expression within delaminated cells, suggesting it was lost once cells became mis-specified, or that they had already progressed within the apoptotic pathway), our observations suggest that some of these cells derive from the anterior or pharyngeal endoderm, which might be primed to undergo programmed cell death. Unfortunately, it was beyond the scope of this work to perform TUNEL staining in our SP600125 or SP600125 + Z-VAD-fmk rescue experiments, but we predict that the pattern of cell death would differ relative to controls specifically within anterior ventral endoderm. Thus, JNK inhibition may not directly activate Caspase-3-mediated apoptosis within amphioxus tissues proper, which is instead induced secondarily within "shed" cells. An alternative, which we do not favour, is that cCasp-3 immunoreactivity is simply not sensitive enough in early stages of delamination to detect in our experiments.

Regardless of how cells are lost from tissues, changes in the levels or patterns of proliferation relative to baseline might in part account for the phenotypes we observe. Under control conditions, we saw broad but variable patterns of cell division, with relatively large numbers of PH3^{+} cells in the endoderm throughout development. Few cells were labelled in the nerve cord, somites, or notochord in any one individual, with more consistent proliferation in the posterior notochord and nerve cord within the tailbud. Our results are broadly congruent with BrDU pulse-chase experiments (Holland and Holland, 2006), which label dividing cells and their progeny, demonstrating that much of the raw material of the notochord is generated by proliferation in the gastrula and early neurula, with proliferation gradually declining at later stages except in the anterior and posterior tips. Although it was too challenging to accurately quantify proliferation in treated embryos here, these patterns appear to be maintained after exposure to SP600125, and therefore are unlikely to account for shortening of the primary axis. However, we cannot discount the possibility that JNK inhibition might in fact delay or arrest passage through the cell cycle, and thus contribute to defects in elongation and patterning. For instance, treatment with SP600125 prolongs mitotic progression through S and G2 phases of the cell cycle in synchronised mouse cells; such growth-inhibited cells eventually die *via* apoptosis (Du et al., 2004). SP600125 also enhances Caspase-3 dependent apoptosis in tumour cells (Kuntzen et al., 2005) and prevents transition from the G2 to M phase (Kuntzen et al., 2005; Kim et al., 2010). However, this arrest may occur *independently* of its ability to inhibit JNK, since specific knockdown of *JNK1/2* by siRNA has no effect on PH3 protein levels in cells (Kim et al., 2010). These studies might also help to explain some apparently contradictory results in flatworms, which show both that SP600125 causes G2/M arrest (Tasaki et al., 2011) and faster entry of neoblasts into mitosis after *JNK(RNAi)* (Almuedo-Castillo et al., 2014). Our static snapshots using PH3 immunostaining, which only labels cells in G2/Mitosis, clearly do not capture the complexity of cell cycle dynamics. Nevertheless, even if there is some non-specific effect of SP600125 on proliferation kinetics in our experiments, or a general delay in cell division, we consider it unlikely that a decrease in proliferation alone can account for our results.

In light of these observations, we currently favour the hypothesis that inhibiting JNK signalling tips the balance in favour of cellular loss, resulting in shortened and mis-proportioned embryos, with cell division unable to compensate; apoptosis is effectively a secondary effect of a loss of cellular polarity and tissue integrity. However, careful analysis of cell cycle dynamics and cellular behaviours of labelled cells *in vivo*, as well as more detailed studies on JNK-mediated cell death in specific tissues using molecular markers, will be required to test and refine this model to dissect the context-dependency of apoptosis and proliferation during amphioxus morphogenesis.

Phosphorylated ERK Activation May Be a Mechanism for Cellular Extrusion in Notochord and Neural Cells

Our observation that JNK inhibition appears to cause cytoskeletal changes, rounding, and delamination of cells into extracellular spaces, combined with CE-type defects, might suggest that active cellular extrusion of mis-specified cells is occurring, followed by (or concomitant with) their Caspase-3-mediated apoptosis. Both live and apoptotic cellular extrusion are important regulators of growth and cell density and are critical for maintaining tissue homeostasis (Ohsawa et al., 2018). The orientation of cellular extrusion – either basally or apically depending on the model – can lead to either hypertrophic overgrowth (cancer) or apoptosis. Importantly, epithelial extrusion is a highly dynamic process *in vivo*, requiring coordination of the actin cytoskeleton and cell junction dis/assembly in cells fated to die as well as their neighbours to maintain tissue integrity (reviewed in Gagliardi and Primo, 2019). Intriguingly, mechanical stress due to crowding alone within a proliferating tissue is also sufficient to cause Caspase induction and basal delamination in the fly notum and may depend on the direction of cell division (Levayer et al., 2016; reviewed in Tada, 2021). Thus, various types of signal can lead to cell extrusion when cells are stressed or dying.

The role of JNK in these processes is complex, and results are often contradictory depending on the cellular context. Typically, inappropriate JNK induction in "loser cells" is associated with their cellular elimination and apoptosis in a cell competition scenario, but can also result in hyperproliferation, for instance during regeneration (Pinal et al., 2019). It has also been suggested that cell cycle progression in G2/M is a checkpoint for cellular extrusion, such that "unfit" cells die if they are in G2 during cell competition (Tada, 2021). In the context of our work in amphioxus, SP600125 treatment might lead to extrusion of cells in several ways, including changes in apicobasal polarity or orientation of cell division, cell crowding, and competition. These might be particularly prevalent in tissues undergoing active cellular rearrangement and extension, such as the notochord at mid-neurula stages, or that proliferate and remodel extensively, such as the anterior ventral endoderm.

One of the key findings of this study was that after JNK inhibition, ERK signalling is activated in the neighbours of cells undergoing extrusion and dying, as evidenced by instances of nuclear fragmentation and Caspase-3 cleavage. This phenomenon was most apparent in the monolayer of the early notochord, but pERK^{+} cells were also sometimes seen in the nerve cord. Our data are both consistent with and appear to contradict results in other systems. For instance, in zebrafish, correct regulation of ERK activity is important for axis elongation and notochord differentiation from chordamesoderm, but not for its early patterning or specification (Hawkins et al., 2008), similarly to amphioxus and contrary to its early role in notochord determination in *Ciona* (Minokawa et al., 2001; Sakabe et al., 2006). U0126 treatment (which blocks ERK1/2 phosphorylation and thus ERK activation), however, results in an increase in TUNEL^{+} apoptotic cells in 24hpf zebrafish, but not in the notochord (Hawkins et al., 2008), which is inconsistent with our own results. In a similar vein, during *Ciona* tail regression, ERK activity induces apoptosis in the notochord and epidermis (Chambon et al., 2002, 2007). However, SP600125 or U0126 treatment both reduce the number of TUNEL^{+} cells in the tail. Chambon et al. (2007) propose a model whereby normal JNK activation in the CNS induces apoptosis through activation of ERK in adjacent tissues, including the notochord, *via* modification of the ECM. The endoderm and CNS themselves are non-receptive and escape the death signals (Chambon et al., 2002, 2007). We would argue that some of these apparent discrepancies across studies may be due to regulative vs. determined developmental mode, the level of cellular resolution, variable tissue-specific and temporal responses, or differences between ERK and JNK targets. It is of further note that in the *Drosophila* epidermis, mis-specified cells exhibit aberrant EGFR signalling with ERK activation, leading to apoptosis, whereas endogenous ERK activation apparently promotes cell survival (Crossman et al., 2018), and that cumulative ERK dosage impacts developmental income (Johnson and Toettcher, 2019). Thus, ERK response varies both across physiological conditions as well as under different aberrant ones.

Several lines of evidence suggest that regulation of ERK might also link cellular extrusion, cell competition, and cellular survival outcomes with mechanical forces within tissues (Moreno et al., 2019; Valon et al., 2021). Compaction and/or a reduction in tension results in down-regulation of EGFR/ERK signalling in looser cells, leading to their elimination *via* apoptosis; conversely local tissue stretching was correlated with activation of ERK (Moreno et al., 2019). Interestingly, compression in the *Xenopus* embryonic epithelium by high gravity causes the cells to activate FGFR/ERK signalling by sensing tensile forces at the adherens and tight junctions (Kinoshita et al., 2020). While we did not directly manipulate forces, it seems likely that SP600125-treated embryos are subjected to considerable compressive forces and tension due to incompatible changes in growth and cellular behaviours across different tissues during neurulation and gastrulation. These conflicts might be expected to be strongest in the notochord as it fails to elongate properly and cells compete for space, and in the anterior somites as they continue to be compressed by cellular accumulation. Punctate β-catenin accumulation at the membranes around extruding (normally pERK low) cells might provide some limited evidence to support an effect on the adherens junctions in the notochord. The fact we saw no convincing pERK modulation in anterior endoderm cells, which appear to also round up and delaminate, can be explained if the effect of JNK inhibition here were due to changes in the plane of division symmetry during proliferation or the cell cycle, rather than from mechanical stresses associated with the highly regulated requirements for CE seen in notochord. Although few specific antibodies are available in cephalochordates, it would be interesting to further examine the effects of Wnt/JNK and other pathways known to affect cellular movements or polarity (e.g., ROCK) on both ERK activity and distribution of apicobasal membrane proteins.

Perspectives

This research presents the first data supporting a putative role for Wnt/JNK signalling in cell behaviour and cell fate decisions during cephalochordate development. Specifically, treatment with the SP600125 inhibitor causes axis shortening consistent with CE defects, accompanied by dramatic changes in the notochord both at the cellular and the molecular levels. Anterior endoderm patterning and somite derivatives are also affected by inhibition of JNK, hinting at some conserved mechanisms with vertebrates, in addition to more ancient roles in regulating the balance between apoptosis and cell survival. Combining the benefits of invertebrate (e.g., a single orthologue) and vertebrate models (chordate body plan and morphogenesis), amphioxus is therefore a good system for studying the myriad and complex roles of JNK signalling in cell biology and morphogenesis.

Our work has several limitations, including the reliance on fixed material and therefore static snapshots of highly dynamic processes *in vivo*. It remains difficult to effectively immobilise live amphioxus embryos, which at all stages are extremely fragile. The ciliary movement embryos use to glide through the water column, even prior to active muscle contraction-dependent swimming in larval stages, has proven difficult to pharmacologically inhibit without causing other defects. Even trying to hold embryos in place in agarose, without applying undue pressure using coverslips – which can itself modulate normal cellular behaviours and even gene expression – is short lived as the embryos ultimately rotate in place. In addition, although the reliance on pharmacological treatment presents notable advantages, it does prevent the dissection of the molecular mechanisms underlying the context-dependent functions of JNK achievable with more targeted genetic approaches. However, the generation of transgenic or mutant lines (e.g., Li et al., 2017), for example to allow temporal and spatial analysis of cell movements at high resolution, is still limited to a handful of laboratories with access to tropical species with short life cycles, and the resources to maintain breeding animal colonies long term. These considerations apply to many non-traditional, particularly marine model systems. Nevertheless, advances such as hydrogel encapsulation (Burnett et al., 2018) or acoustic trapping combined with live imaging (Yang et al., 2019) have shown proof of principle that embryos can be maintained alive and develop normally to allow studies of cellular movements. Combined with better fluorescent dyes, for instance to label cell membranes, and fast 3D imaging, as well as further developments in cell tracking, imaging dynamic cells movements in amphioxus should soon prove more tractable, permitting more detailed analyses of the evolution of morphogenetic and cellular processes at the origin of chordates.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

IMLS conceived the study, performed the experiments, analysed the data, provided resources, and wrote the manuscript. MTE modelled SP600125 and JNK protein binding and contributed to **Figure 1**. HE and JG-F provided resources. All authors read and approved the final manuscript.

FUNDING

HE was supported by the Centre National de la Recherche Scientifique and Agence Nationale de la Recherche (ANR) grants nos. ANR-16-CE12-0008-01 and ANR-19-CE13-0011. JG-F was supported by BFU2017-861152-P and PID2020-117820GB-100 grants from the Ministerio de Ciencia y Innovación (Spain). This study was in part supported by the European community through a Marie Curie fellowship (FP7-People-IEF-2008, Project 236867) to IMLS. This research was funded in part by the Wellcome Trust grant ISSF3 (grant number 204821/Z/16/Z) to IMLS. For the purpose of open access, the author has applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission.

ACKNOWLEDGMENTS

IMLS would like to thank Marcus Bischoff for stimulating discussion about cellular extrusion.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.749806/full#supplementary-material

Supplementary Figure 1 | Alignment of three JNK sequences from human (MAPK8, 9, and 10) and single orthologues from amphioxus species *Branchiostoma floridae*, *Branchiostoma lanceolatum*, and *Branchiostoma belcheri*. Transcriptomic data were used to manually correct errors in *B. lanceolatum* gene model BL14252 from Ensembl to generate a consensus sequence. The alignment was generated in Seaview using Muscle.

Supplementary Figure 2 | Defects in invagination and developmental arrest during gastrulation after SP600125 treatment (5 μM) at blastula stage. Exogastrulae arrested during invagination by JNK inhibition are boxed in yellow. Column 1 embryos were stopped at the gastrula stage (control to left). Lateral (top) and blastoporal (bottom) views are shown for each marker gene. Columns 2 and 3 embryos were stopped at neurula stages (control to right). Lateral (top) and dorsal (bottom) views are shown for each marker gene. In some cases blastoporal views are shown (*dll*, *nodal*, and *neurogenin*). Wild type gastrula stages are shown on the left for comparison to show how gene expression resembles earlier stages of development. For *dkk1/2/4*, treated embryos allowed to further develop are boxed in cyan, showing no change in gene expression relative to those stopped earlier. Scale bars = 50 μM.

Supplementary Figure 3 | Additional markers used to characterise the *orca* phenotype show inconclusive results or no change in expression relative to controls. Embryos were treated with SP600125 as per **Figure 2**. Lateral views show anterior to the left and dorsal up. Dorsal views show anterior to the left. Dorsal views for *otx* are omitted as they were uninformative (the underlying endodermal expression was too strong, masking the neural domains). The black outlined arrowhead highlights loss of separation between neural plate and endoderm expression domains of *otx* in treated embryos. The white asterisk in *wnt4* shows non-specific staining caused by particulate material stuck to the embryo surface. Generally, embryos were stained longer than required to ensure all real domains became revealed, resulting in some general background (e.g., bluish/purplish tinge, especially in mesendoderm, in both treated and controls). Scale bars = 50 μM.

Supplementary Figure 4 | Limited apoptosis in late embryos after weak JNK inhibition. **(A,A′,A″)** T0 stage DMSO control embryo (stitching of two separate confocal images). Perforation of the pharynx is imminent (yellow asterisks) but little cleaved Caspase-3 is evident. **(B,B′,B″)** Stage matched 2.5 μM SP600125 treated embryos have a disorganised but differentiated central notochord and lack anterior structures, including the pharynx (yellow asterisks), which consists of many rounding cells ventrally (blue arrowheads). The gut contains cleaved Caspase-3-positive pycnotic cellular material (**B″**, insets x, x′ and y, y′), and occasional notochord cells are apoptotic and in the process of extrusion (white arrowheads in insets x, x′ and y, y′). Overall, however, there is little evidence of Caspase-3-dependent apoptosis at these late stages. Phalloidin staining of F-Actin is in blue, cleaved Caspase-3 (cCasp-3) immunoreactivity is in magenta, and DAPI stained nuclei are in grey. nt, neural tube; cv, cerebral vesicle; no, notochord; ph, pharynx. Embryos are orientated with anterior to the left and dorsal up. Scale bars = 50 μM.

Supplementary File 1 | Sequences of new probes used in this study.

REFERENCES

Aikin, T. J., Peterson, A. F., Pokrass, M. J., Clark, H. R., and Regot, S. (2020). MAPK activity dynamics regulate non-cell autonomous effects of oncogene expression. *Elife* 9:e60541. doi: 10.7554/eLife.60541.sa2

Ait-Hamlat, A., Zea, D. J., Labeeuw, A., Polit, L., Richard, H., and Laine, E. (2020). Transcripts' evolutionary history and structural dynamics give mechanistic insights into the functional diversity of the JNK family. *J. Mol. Biol.* 432, 2121–2140. doi: 10.1016/j.jmb.2020.01.032

Albuixech-Crespo, B., López-Blanch, L., Burguera, D., Maeso, I., Sánchez-Arrones, L., Moreno-Bravo, J. A., et al. (2017). Molecular regionalization of the developing amphioxus neural tube challenges major partitions of the vertebrate brain. *PLoS Biol.* 15:e2001573. doi: 10.1371/journal.pbio.2001573

Aldea, D., Subirana, L., Keime, C., Meister, L., Maeso, I., Marcellini, S., et al. (2019). Genetic regulation of amphioxus somitogenesis informs the evolution of the vertebrate head mesoderm. *Nat. Ecol. Evol.* 3, 1233–1240. doi: 10.1038/s41559-019-0933-z

Almuedo-Castillo, M., Crespo, X., Seebeck, F., Bartscherer, K., Salò, E., and Adell, T. (2014). JNK controls the onset of mitosis in planarian stem cells and triggers apoptotic cell death required for regeneration and remodeling. *PLoS Genet.* 10:e1004400. doi: 10.1371/journal.pgen.1004400

Babonis, L. S., and Martindale, M. Q. (2017). Phylogenetic evidence for the modular evolution of metazoan signalling pathways. *Phil. Trans. Royal Soc. B: Biol. Sci.* 372, 20150477. doi: 10.1098/rstb.2015.0477

Barton-Owen, T. B., Ferrier, D. E., and Somorjai, I. M. (2018). Pax3/7 duplicated and diverged independently in amphioxus, the basal chordate lineage. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-27700-x

Bayascas, J., Yuste, V., Benito, E., Garcia-Fernàndez, J., and Comella, J. X. (2002). Isolation of AmphiCASP-3/7, an ancestral caspase from amphioxus (*Branchiostoma floridae*). Evolutionary considerations for vertebrate caspases. *Cell Death Differ.* 9, 1078–1089. doi: 10.1038/sj.cdd.4401075

Bennett, B. L., Sasaki, D. T., Murray, B. W., O'Leary, E. C., Sakata, S. T., Xu, W., et al. (2001). SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. *Proc. Nat. Acad. Sci. U.S.A.* 98, 13681–13686. doi: 10.1073/pnas.251194298

Bertrand, S., and Escrivà, H. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* 138, 4819–4830. doi: 10.1242/dev.066720

Bertrand, S., Camasses, A., Somorjai, I., Belgacem, M. R., and Chabrol, O. (2011). Amphioxus FGF signaling predicts the acquisition of vertebrate morphological traits. *Proc. Nat. Acad. Sci. U.S.A.* 108, 9160–9165. doi: 10.1073/pnas.1014235108

Bertrand, S., Campo-Paysaa, F., Camasses, A., García-Fernàndez, J., and Escrivà, H. (2009). Actors of the tyrosine kinase receptor downstream signaling pathways in amphioxus. *Evol. Dev.* 11, 13–26. doi: 10.1111/j.1525-142x.2008.00299.x

Bikkavilli, R. K., and Malbon, C. C. (2009). Mitogen-activated protein kinases and Wnt/β-catenin signaling: molecular conversations among signaling pathways. *Commun. Integr. Biol.* 2, 46–49. doi: 10.4161/cib.2.1.7503

Bosze, B., Ono, Y., Mattes, B., Sinner, C., Gourain, V., Thumberger, T., et al. (2020). Pcdh18a regulates endocytosis of E-cadherin during axial mesoderm development in zebrafish. *Histochem. Cell. Biol.* 154, 463–480. doi: 10.1007/s00418-020-01887-5

Bozzo, M., Lacalli, T. C., Obino, V., Caicci, F., Marcenaro, E., Bachetti, T., et al. (2021). Amphioxus neuroglia: molecular characterization and evidence for early compartmentalization of the developing nerve cord. *Glia* 69, 1654–1678. doi: 10.1002/glia.23982

Brunt, L., Greicius, G., Rogers, S., Evans, B. D., Virshup, D. M., Wedgwood, K. C. A., et al. (2021). Vangl2 promotes the formation of long cytonemes to enable distant Wnt/β-catenin signaling. *Nat. Commun.* 12:2058. doi: 10.1038/s41467-021-22393-9

Burnett, K., Edsinger, E., and Albrecht, D. R. (2018). Rapid and gentle hydrogel encapsulation of living organisms enables long-term microscopy over multiple hours. *Commun. Biol.* 1, 1–10. doi: 10.1038/s42003-018-0079-6

Cai, C., Lin, J., Sun, S., and He, Y. (2016). JNK inhibition inhibits lateral line neuromast hair cell development. *Front. Cell. Neurosci.* 10:19. doi: 10.3389/fncel.2016.00019

Candiani, S., Holland, N. D., Oliveri, D., Parodi, M., and Pestarino, M. (2008). Expression of the amphioxus Pit-1 gene (AmphiPOU1F1/Pit-1) exclusively in the developing preoral organ, a putative homolog of the vertebrate adenohypophysis. *Brain Res. Bull.* 75, 324–330. doi: 10.1016/j.brainresbull.2007.10.023

Canovas, B., and Nebreda, A. R. (2021). Diversity and versatility of p38 kinase signalling in health and disease. *Nat. Rev. Mol. Cell Biol.* 22, 346–366. doi: 10.1038/s41580-020-00322-w

Carvalho, J. E., Lahaye, F., Yong, L. W., Croce, J. C., Escrivá, H., Yu, J. K., et al. (2021). An updated staging system for cephalochordate development: one table suits them all. *Front. Cell Dev. Biol.* 9:668006. doi: 10.3389/fcell.2021.668006

Castro-Torres, R. D., Busquets, O., Parcerisas, A., Verdaguer, E., Olloquequi, J., Ettcheto, M., et al. (2020). Involvement of JNK1 in neuronal polarization during brain development. *Cells* 9:1897. doi: 10.3390/cells9081897

Chambon, J. P., Nakayama, A., Takamura, K., McDougall, A., and Satoh, N. (2007). ERK-and JNK-signalling regulate gene networks that stimulate metamorphosis and apoptosis in tail tissues of ascidian tadpoles. *Development* 134, 1203–1219. doi: 10.1242/dev.002220

Chambon, J. P., Soule, J., Pomies, P., Fort, P., Sahuquet, A., Alexandre, D., et al. (2002). Tail regression in *Ciona* intestinalis (Prochordate) involves a Caspase-dependent apoptosis event associated with ERK activation. *Development* 129, 3105–3114. doi: 10.1242/dev.129.13.3105

Crossman, S. H., Streichan, S. J., and Vincent, J. P. (2018). EGFR signaling coordinates patterning with cell survival during *Drosophila* epidermal development. *PLoS Biol.* 16:e3000027. doi: 10.1371/journal.pbio.3000027 doi: 10.1371/journal.pbio.3000027

Dailey, S. C., Febrero Planas, R., Rossell Espier, A., Garcia-Fernandez, J., and Somorjai, I. M. (2016). Asymmetric distribution of pl10 and bruno2, new members of a conserved core of early germline determinants in cephalochordates. *Front. Ecol. Evol.* 3:156. doi: 10.3389/fevo.2015.00156

Dailey, S. C., Kozmikova, I., and Somorjai, I. (2017). Amphioxus Sp5 is a member of a conserved Specificity Protein complement and is modulated by Wnt/β-catenin signalling. *Int. J. Dev. Biol.* 61, 723–732. doi: 10.1387/ijdb.170205is

D'Aniello, S., Irimia, M., Maeso, I., Pascual-Anaya, J., Jiménez-Delgado, S., Bertrand, S., et al. (2008). Gene expansion and retention leads to a diverse tyrosine kinase superfamily in amphioxus. *Mol. Biol. Evol.* 25, 1841–1854. doi: 10.1093/molbev/msn132 doi: 10.1093/molbev/msn132

Dehal, P., and Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:e314. doi: 10.1371/journal.pbio.0030314

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* 19, 4008. doi: 10.1093/nar/19.14.4008

Du, L., Lyle, C. S., Obey, T. B., Gaarde, W. A., Muir, J. A., Bennett, B. L., et al. (2004). Inhibition of cell proliferation and cell cycle progression by specific inhibition of basal JNK activity: evidence that mitotic Bcl-2 phosphorylation is JNK-independent. *J. Biol. Chem.* 279, 11957–11966. doi: 10.1074/jbc.m304935200

Dush, M. K., and Nascone-Yoder, N. M. (2013). Jun N-terminal kinase maintains tissue integrity during cell rearrangement in the gut. *Development* 140, 1457–1466. doi: 10.1242/dev.086850

Escrivà, H. (2018). My favorite animal, amphioxus: unparalleled for studying early vertebrate evolution. *Bioessays* 40:e1800130. doi: 10.1002/bies.201800130

Fuentes, M., Benito, E., Bertrand, S., Paris, M., Mignardot, A., Godoy, L., et al. (2007). Insights into spawning behavior and development of the European amphioxus (*Branchiostoma lanceolatum*). *J. Exp. Zool. B: Mol. Dev. Evol.* 308, 484–493.

Gagliardi, P. A., and Primo, L. (2019). Death for life: a path from apoptotic signaling to tissue-scale effects of apoptotic epithelial extrusion. *Cell. Mol. Life Sci.* 76, 3571–3581. doi: 10.1007/s00018-019-03153-x

Gao, C., and Chen, Y. G. (2010). Dishevelled: the hub of Wnt signaling. *Cell Signal.* 22, 717–727. doi: 10.1016/j.cellsig.2009.11.021

Glasauer, S. M. K., and Neuhauss, S. C. F. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol. Genet. Genomics* 289, 1045–1060. doi: 10.1007/s00438-014-0889-2

Gros, J., Hu, J. K. H., Vinegoni, C., Feruglio, P. F., Weissleder, R., and Tabin, C. J. (2010). WNT5A/JNK and FGF/MAPK pathways regulate the cellular events shaping the vertebrate limb bud. *Curr. Biol.* 20, 1993–2002. doi: 10.1016/j.cub.2010.09.063

Guerin, D. J., Kha, C. X., and Tseng, K. A. S. (2021). From cell death to regeneration: rebuilding after injury. *Front. Cell Dev. Biol.* 9:547. doi: 10.3389/fcell.2021.655048

Hardy, K. M., Garriock, R. J., Yatskievych, T. A., D'Agostino, S. L., Antin, P. B., and Krieg, P. A. (2008). Non-canonical Wnt signaling through Wnt5a/b and a novel Wnt11 gene, Wnt11b, regulates cell migration during avian gastrulation. *Dev. Biol.* 320, 391–401. doi: 10.1016/j.ydbio.2008.05.546 doi: 10.1016/j.ydbio.2008.05.546

Hawkins, T. A., Cavodeassi, F., Erdélyi, F., Szabó, G., and Lele, Z. (2008). The small molecule Mek1/2 inhibitor U0126 disrupts the chordamesoderm to notochord transition in zebrafish. *BMC Dev. Biol.* 8:42. doi: 10.1186/1471-213X-8-42

He, Y., Cai, C., Sun, S., Wang, X., Li, W., and Li, H. (2016). Effect of JNK inhibitor SP600125 on hair cell regeneration in zebrafish (*Danio rerio*) larvae. *Oncotarget* 7, 51640–51650. doi: 10.18632/oncotarget.10540

Heisenberg, C. P., Tada, M., Rauch, G. J., Saúde, L., Concha, M. L., Geisler, R., et al. (2000). Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. *Nature* 405, 76–81. doi: 10.1038/35011068

Holland, N. D. (2018). Formation of the initial kidney and mouth opening in larval amphioxus studied with serial blockface scanning electron microscopy (SBSEM). *EvoDevo* 9:16.

Holland, N. D., and Holland, L. Z. (2006). Stage- and tissue-specific patterns of cell division in embryonic and larval tissues of amphioxus during normal development. *Evol. Dev.* 8, 142–149. doi: 10.1111/j.1525-142x.2006.00085.x

Holland, N. D., Paris, M., and Koop, D. (2009). The club-shaped gland of amphioxus: export of secretion to the pharynx in pre-metamorphic larvae and apoptosis during metamorphosis. *Acta Zool.* 90, 372–379. doi: 10.1111/j.1463-6395.2008.00379.x

Houssin, N. S., Bharathan, N. K., Turner, S. D., and Dickinson, A. J. (2017). Role of JNK during buccopharyngeal membrane perforation, the last step of embryonic mouth formation. *Dev. Dyn.* 246, 100–115. doi: 10.1002/dvdy.24470

Irimia, M., Denuc, A., Burguera, D., Somorjai, I., Martín-Durán, J. M., Genikhovich, G., et al. (2011). Stepwise assembly of the Nova-regulated alternative splicing network in the vertebrate brain. *Proc. Nat. Acad. Sci. U.S.A.* 108, 5319–5324. doi: 10.1073/pnas.1012333108

Johnson, H. E., and Toettcher, J. E. (2019). Signaling dynamics control cell fate in the early *Drosophila embryo*. *Dev. Cell.* 48, 361–370. doi: 10.1016/j.devcel.2019.01.009

Kalapos, B., Hlavová, M., Nádai, T. V., Galiba, G., Bišová, K., Dóczi, R., et al. (2019). Early evolution of the Mitogen-activated protein kinase family in the plant kingdom. *Sci. Rep.* 9:4094. doi: 10.1038/s41598-019-40751-y

Karimi, K., Fortriede, J. D., Lotay, V. S., Burns, K. A., Wang, D. Z., Fisher, M. E., et al. (2018). Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res.* 46, D861–D868. doi: 10.1093/nar/gkx936

Keller, R., and Sutherland, A. (2020). Convergent extension in the amphibian, *Xenopus laevis*. *Curr. Top. Dev. Biol.* 136, 271–317. doi: 10.1016/bs.ctdb.2019.11.013

Kim, E. K., and Choi, E. J. (2010). Pathological roles of MAPK signaling pathways in human diseases. *Biochim. Biophys. Acta* 1802, 396–405. doi: 10.1016/j.bbadis.2009.12.009

Kim, G. H., and Han, J. K. (2005). JNK and ROKα function in the noncanonical Wnt/RhoA signaling pathway to regulate *Xenopus* convergent extension movements. *Dev. Dyn.* 232, 958–968. doi: 10.1002/dvdy.20262

Kim, J., Lee, J., Margolis, R., and Fotedar, R. (2010). SP600125 suppresses Cdk1 and induces endoreplication directly from G2 phase, independent of JNK inhibition. *Oncogene* 29, 1702–1716. doi: 10.1038/onc.2009.464

Kinoshita, N., Hashimoto, Y., Yasue, N., Suzuki, M., Cristea, I. M., and Ueno, N. (2020). Mechanical stress regulates epithelial tissue integrity and stiffness through the FGFR/Erk2 signaling pathway during embryogenesis. *Cell Rep.* 30, 3875–3888.e3. doi: 10.1016/j.celrep.2020.02.074 doi: 10.1016/j.celrep.2020.02.074

Koop, D., Holland, L. Z., Setiamarga, D., Schubert, M., and Holland, N. D. (2011). Tail regression induced by elevated retinoic acid signaling in amphioxus larvae occurs by tissue remodeling, not cell death. *Evol. Dev.* 13, 427–435. doi: 10.1111/j.1525-142X.2011.00501.x

Kozmikova, I., and Kozmik, Z. (2020). Wnt/β-catenin signaling is an evolutionarily conserved determinant of chordate dorsal organizer. *Elife* 9:e56817. doi: 10.7554/eLife.56817.sa2

Krasovec, G., Robine, K., Quéinnec, E., Karaiskou, A., and Chambon, J. P. (2019). Ci-hox12 tail gradient precedes and participates in the control of the apoptotic-dependent tail regression during *Ciona larva* metamorphosis. *Dev. Biol.* 448, 237–246. doi: 10.1016/j.ydbio.2018.12.010

Krens, S. G., Spaink, H. P., and Snaar-Jagalska, B. E. (2006). Functions of the MAPK family in vertebrate-development. *FEBS Lett.* 580, 4984–4990. doi: 10.1016/j.febslet.2006.08.025

Kuan, C. Y., Yang, D. D., Roy, D. R. S., Davis, R. J., Rakic, P., and Flavell, R. A. (1999). The Jnk1 and Jnk2 protein kinases are required for regional specific apoptosis during early brain development. *Neuron* 22, 667–676. doi: 10.1016/s0896-6273(00)80727-8 doi: 10.1016/s0896-6273(00)80727-8

Kuntzen, C., Sonuc, N., De Toni, E. N., Opelz, C., Mucha, S. R., Gerbes, A. L., et al. (2005). Inhibition of c-Jun-N-terminal-kinase sensitizes tumor cells to CD95-induced apoptosis and induces G2/M cell cycle arrest. *Cancer Res.* 65, 6780–6788. doi: 10.1158/0008-5472.CAN-04-2618

Levayer, R., Dupont, C., and Moreno, E. (2016). Tissue crowding induces caspase-dependent competition for space. *Curr. Biol.* 26, 670–677. doi: 10.1016/j.cub.2015.12.072 doi: 10.1016/j.cub.2015.12.072

La Marca, J. E., and Richardson, H. E. (2020). Two-faced: Roles of JNK signalling during tumourigenesis in the *Drosophila* model. *Front. Cell Dev. Biol.* 8:42. doi: 10.3389/fcell.2020.00042

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/S0021889892009944

Lavoie, H., Gagnon, J., and Therrien, M. (2020). ERK signalling: a master regulator of cell behaviour, life and fate. *Nat. Rev. Mol. Cell. Biol.* 21, 607–632. doi: 10.1038/s41580-020-0255-7

Lee, M. H., Koria, P., Qu, J., and Andreadis, S. T. (2009). JNK phosphorylates beta-catenin and regulates adherens junctions. *FASEB J.* 23, 3874–3883. doi: 10.1096/fj.08-117804

Letunic, I., Copley, R. R., and Bork, P. (2002). Common exon duplication in animals and its role in alternative splicing. *Hum. Mol. Genet.* 11, 1561–1567. doi: 10.1093/hmg/11.13.1561

Li, G., Liu, X., Xing, C., Zhang, H., Shimeld, S. M., and Wang, Y. (2017). Cerberus–Nodal–Lefty–Pitx signaling cascade controls left–right asymmetry in amphioxus. *Proc. Nat. Acad. Sci. U.S.A.* 114, 3684–3689. doi: 10.1073/pnas.1620519114

Li, L., Ning, G., Yang, S., Yan, Y., Cao, Y., and Wang, Q. (2019). BMP signaling is required for nkx2. 3-positive pharyngeal pouch progenitor specification in zebrafish. *PLoS Genet.* 15:e1007996. doi: 10.1371/journal.pgen.1007996

Lin, C. Y., Lu, M. Y. J., Yue, J. X., Li, K. L., Le Pétillon, Y., Yong, L. W., et al. (2020). Molecular asymmetry in the cephalochordate embryo revealed by single-blastomere transcriptome profiling. *PLoS Genet.* 16:e1009294. doi: 10.1371/journal.pgen.1009294

Long, J. T., Irwin, L., Enomoto, A. C., Grow, Z., Ranck, J., and Peeler, M. T. (2015). Jun N-terminal kinase activity is required for invagination but not differentiation of the sea urchin archenteron. *Genesis* 53, 762–769. doi: 10.1002/dvg.22898

Marlétaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564, 64–70. doi: 10.1038/s41586-018-0734-6

Martín-Blanco, E., Pastor-Pareja, J. C., and García-Bellido, A. (2000). JNK and decapentaplegic signaling control adhesiveness and cytoskeleton dynamics during thorax closure in Drosophila. *Proc. Nat. Acad. Sci. U.S.A.* 97, 7888–7893. doi: 10.1073/pnas.97.14.7888

Mazzotta, S., Neves, C., Bonner, R. J., Bernardo, A. S., Docherty, K., and Hoppler, S. (2016). Distinctive roles of canonical and noncanonical Wnt signaling in human embryonic cardiomyocyte development. *Stem Cell Rep.* 7, 764–776. doi: 10.1016/j.stemcr.2016.08.008

Minokawa, T., Yagi, K., Makabe, K. W., and Nishida, H. (2001). Binary specification of nerve cord and notochord cell fates in ascidian embryos. *Development* 128, 2007–2017. doi: 10.1242/dev.128.11.2007

Moreno, E., Valon, L., Levillayer, F., and Levayer, R. (2019). Competition for space induces cell elimination through compaction-driven ERK downregulation. *Curr. Biol.* 29, 23–34.e8. doi: 10.1016/j.cub.2018.11.007 doi: 10.1016/j.cub.2018.11.007

Ohsawa, S., Vaughen, J., and Igaki, T. (2018). Cell extrusion: A stress-responsive force for good or evil in epithelial homeostasis. *Dev. Cell.* 44, 284–296. doi: 10.1016/j.devcel.2018.01.009

Pai, Y. J., Abdullah, N. L., Mohd-Zin, S. W., Mohammed, R. S., Rolo, A., Greene, N. D., et al. (2012). Epithelial fusion during neural tube morphogenesis. *Birth Defects Res. A* 94, 817–823. doi: 10.1002/bdra.23072

Pinal, N., Calleja, M., and Morata, G. (2019). Pro-apoptotic and pro-proliferation functions of the JNK pathway of Drosophila: roles in cell competition, tumorigenesis and regeneration. *Open Biol.* 9:180256. doi: 10.1098/rsob.180256

Ricci, L., and Srivastava, M. (2018). Wound-induced cell proliferation during animal regeneration. *Wiley Interdiscip. Rev. Dev. Biol.* 7, e321. doi: 10.1002/wdev.321

Ríos-Barrera, L. D., and Riesgo-Escovar, J. R. (2013). Regulating cell morphogenesis: the *Drosophila* Jun N-terminal kinase pathway. *Genesis* 51, 147–162. doi: 10.1002/dvg.22354

Sakabe, E., Tanaka, N., Shimozono, N., Gojobori, T., and Fujiwara, S. (2006). Effects of U0126 and fibroblast growth factor on gene expression profile in *Ciona intestinalis* embryos as revealed by microarray analysis. *Dev. Growth Differ.* 48, 391–400. doi: 10.1111/j.1440-169X.2006.00877.x

Santos-Ledo, A., Washer, S., Dhanaseelan, T., Eley, L., Alqatani, A., Chrystal, P. W., et al. (2020). Alternative splicing of jnk1a in zebrafish determines first heart field ventricular cardiomyocyte numbers through modulation of hand2 expression. *PLoS Genet.* 16:e1008782. doi: 10.1371/journal.pgen.1008782

Satou, Y., Sasakura, Y., Yamada, L., Imai, K. S., Satoh, N., and Degnan, B. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. V. Genes for receptor tyrosine kinase pathway and Notch signaling pathway. *Dev. Genes Evol.* 213, 254–263. doi: 10.1007/s00427-003-0317-9

Scapin, G., Patel, S. B., Lisnock, J., Becker, J. W., and LoGrasso, P. V. (2003). The structure of JNK3 in complex with small molecule inhibitors: structural basis for potency and selectivity. *Chem. Biol.* 10, 705–712. doi: 10.1016/S1074-5521(03)00159-5

Schellino, R., Boido, M., and Vercelli, A. (2019). JNK signaling pathway involvement in spinal cord neuron development and death. *Cells* 8:1576. doi: 10.3390/cells8121576

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Met.* 9, 676–682. doi: 10.1038/nmeth.2019

Seo, J., Asaoka, Y., Nagai, Y., Hirayama, J., Yamasaki, T., Namae, M., et al. (2010). Negative regulation of wnt11 expression by Jnk signaling during zebrafish gastrulation. *J. Cell. Biochem.* 110, 1022–1037. doi: 10.1002/jcb.22616

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., et al. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538, 336–343. doi: 10.1038/nature19840 doi: 10.1038/nature19840

Shilo, B. Z. (2014). The regulation and functions of MAPK pathways in *Drosophila*. *Methods* 68, 151–159. doi: 10.1016/j.ymeth.2014.01.020

Somorjai, I. M., Escrivà, H., and Garcia-Fernàndez, J. (2012a). Amphioxus makes the cut—again. *Commun. Integr. Biol.* 5, 499–502. doi: 10.4161/cib.21075

Somorjai, I. M., Somorjai, R. L., Garcia-Fernàndez, J., and Escrivà, H. (2012b). Vertebrate-like regeneration in the invertebrate chordate amphioxus. *Proc. Nat. Acad. Sci. U.S.A.* 109, 517–522. doi: 10.1073/pnas.1100045109

Somorjai, I. M., Martí-Solans, J., Diaz-Gracia, M., Nishida, H., Imai, K. S., Escrivà, H., et al. (2018). Wnt evolution and function shuffling in liberal and conservative chordate genomes. *Genome Biol.* 19, 1–17. doi: 10.1186/s13059-018-1468-3

Somorjai, I., Bertrand, S., Camasses, A., Haguenauer, A., and Escriva, H. (2008). Evidence for stasis and not genetic piracy in developmental expression patterns of Branchiostoma lanceolatum and *Branchiostoma floridae*, two amphioxus species that have evolved independently over the course of 200 Myr. *Dev. Genes Evol.* 218, 703–713. doi: 10.1007/s00427-008-0256-6

Stach, T. (1996). On the preoral pit of the larval amphioxus (*Branchiostoma lanceolatum*). *Ann. Sci. Nat. Zool. Biol. Anim.* 17, 129–134.

Steinmetz, P. R., Zelada-Gonzáles, F., Burgtorf, C., Wittbrodt, J., and Arendt, D. (2007). Polychaete trunk neuroectoderm converges and extends by mediolateral cell intercalation. *Proc. Nat. Acad. Sci. U.S.A.* 104, 2727–2732. doi: 10.1073/pnas.0606589104

Tada, M. (2021). The morphogenetic changes that lead to cell extrusion in development and cell competition. *Dev. Biol.* 477, 1–10. doi: 10.1016/j.ydbio.2021.05.003

Tait, S., and Green, D. (2008). Caspase-independent cell death: leaving the set without the final cut. *Oncogene* 27, 6452–6461. doi: 10.1038/onc.2008.311

Tasaki, J., Shibata, N., Sakurai, T., Agata, K., and Umesono, Y. (2011). Role of c-Jun N-terminal kinase activation in blastema formation during planarian regeneration. *Dev. Growth Differ.* 53, 389–400.

Webb, B., and Sali, A. (2016). Comparative protein structure modeling using modeller. *Curr. Protoc. Bioinformatics* 54, 5.6.1–5.6.37. doi: 10.1002/cpbi.3

Valon, L., Davidoviæ, A., Levillayer, F., Villars, A., Chouly, M., Cerqueira-Campos, F., et al. (2021). Robustness of epithelial sealing is an emerging property of local ERK feedback driven by cell elimination. *Dev. Cell* 56, 1700–1711. doi: 10.1016/j.devcel.2021.05.006

Williams, M. L., and Solnica-Krezel, L. (2020). Cellular and molecular mechanisms of convergence and extension in zebrafish. *Curr. Top. Dev. Biol.* 136, 377–407. doi: 10.1016/bs.ctdb.2019.08.001

Xu, C., Liu, R., Zhang, Q., Chen, X., Qian, Y., Fang, W., et al. (2017). The diversification of evolutionarily conserved MAPK cascades correlates with the evolution of fungal species and development of lifestyles. *Genome Biol. Evol.* 9, 311–322. doi: 10.1093/gbe/evw051

Yamanaka, H., Moriguchi, T., Masuyama, N., Kusakabe, M., Hanafusa, H., Takada, R., et al. (2002). JNK functions in the non-canonical Wnt pathway to regulate convergent extension movements in vertebrates. *EMBO Rep.* 3, 69–75.

Yang, Z., Cole, K. L., Qiu, Y., Somorjai, I. M., Wijesinghe, P., Nylk, J., et al. (2019). Light sheet microscopy with acoustic sample confinement. *Nat. Commun.* 10, 1–8. doi: 10.1038/s41467-019-08514-5

Yu, H., Wu, C. L., Wang, X., Ban, Q., Quan, C., Liu, Q., et al. (2019). SP600125 enhances C-2-induced cell death by the switch from autophagy to apoptosis in bladder cancer cells. *J. Exp. Clin. Cancer Res.* 38, 448. doi: 10.1186/s13046-019-1467-6

Yuan, S., Tao, X., Huang, S., Chen, S., and Xu, A. (2014). Comparative immune systems in animals. *Annu. Rev. Anim. Biosci.* 2, 235–258. doi: 10.1146/annurev-animal-031412-103634

Zeke, A., Misheva, M., Reményi, A., and Bogoyevitch, M. A. (2016). JNK signaling: regulation and functions based on complex protein-protein partnerships. *Microbiol. Mol. Biol. Rev.* 80, 793–835. doi: 10.1128/MMBR.00043-14

Zhang, Z., Rankin, S. A., and Zorn, A. M. (2016). Syndecan4 coordinates Wnt/JNK and BMP signaling to regulate foregut progenitor development. *Dev. Biol.* 416, 187–199. doi: 10.1016/j.ydbio.2016.05.025

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Somorjai, Ehebauer, Escrivà and Garcia-Fernàndez. 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) and the copyright owner(s) 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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fcell.2021.749963

Derivedness Index for Estimating Degree of Phenotypic Evolution of Embryos: A Study of Comparative Transcriptomic Analyses of Chordates and Echinoderms

Jason Cheok Kuan Leong[1]*, Yongxin Li[2], Masahiro Uesaka[3], Yui Uchida[1,4], Akihito Omori[5], Meng Hao[2], Wenting Wan[2], Yang Dong[2], Yandong Ren[2], Si Zhang[6], Tao Zeng[6], Fayou Wang[6], Luonan Chen[6,7], Gary Wessel[8], Brian T. Livingston[9], Cynthia Bradham[10], Wen Wang[2,11] and Naoki Irie[1,4]*

[1]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan, [2]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China, [3]RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Japan, [4]Universal Biology Institute, The University of Tokyo, Tokyo, Japan, [5]Sado Island Center for Ecological Sustainability, Niigata University, Niigata, Japan, [6]Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai, China, [7]Key Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou, China, [8]Providence Institute of Molecular Oogenesis, Brown University, Providence, RI, United States, [9]Department of Biological Sciences, California State University, Long Beach, CA, United States, [10]Department of Biology, Boston University, Boston, MA, United States, [11]School of Ecology and Environment, Northwestern Polytechnical University, Xi'an, China

OPEN ACCESS

Edited by:
Stephanie Bertrand,
UMR7232 Biologie Intégrative des Organismes Marins (BIOM), France

Reviewed by:
Cedric Patthey,
Umeå University, Sweden
Ronald Jenner,
Natural History Museum, United Kingdom

***Correspondence:**
Jason Cheok Kuan Leong
jasonleong@bs.s.u-tokyo.ac.jp
Naoki Irie
irie@bs.s.u-tokyo.ac.jp

Specialty section:
This article was submitted to Evolutionary Developmental Biology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 July 2021*
Accepted: *03 November 2021*
Published: *26 November 2021*

Citation:
Leong JCK, Li Y, Uesaka M, Uchida Y, Omori A, Hao M, Wan W, Dong Y, Ren Y, Zhang S, Zeng T, Wang F, Chen L, Wessel G, Livingston BT, Bradham C, Wang W and Irie N (2021) Derivedness Index for Estimating Degree of Phenotypic Evolution of Embryos: A Study of Comparative Transcriptomic Analyses of Chordates and Echinoderms.
Front. Cell Dev. Biol. 9:749963.
doi: 10.3389/fcell.2021.749963

Species retaining ancestral features, such as species called living fossils, are often regarded as less derived than their sister groups, but such discussions are usually based on qualitative enumeration of conserved traits. This approach creates a major barrier, especially when quantifying the degree of phenotypic evolution or degree of derivedness, since it focuses only on commonly shared traits, and newly acquired or lost traits are often overlooked. To provide a potential solution to this problem, especially for inter-species comparison of gene expression profiles, we propose a new method named "derivedness index" to quantify the degree of derivedness. In contrast to the conservation-based approach, which deals with expressions of commonly shared genes among species being compared, the derivedness index also considers those that were potentially lost or duplicated during evolution. By applying our method, we found that the gene expression profiles of penta-radial phases in echinoderm tended to be more highly derived than those of the bilateral phase. However, our results suggest that echinoderms may not have experienced much larger modifications to their developmental systems than chordates, at least at the transcriptomic level. In vertebrates, we found that the mid-embryonic and organogenesis stages were

---

Abbreviations: Aj, sea cucumber (*Apostichopus japonicus*); Cg, oyster (*Crassostrea gigas*); DCO, derivedness-correlative ortholog-groups; Mm, mouse (*Mus musculus*); Gg, chicken (*Gallus gallus*); Ps, Chinese soft-shelled turtle (*Pelodiscus sinensis*); Xl, African clawed frog (*Xenopus laevis*); Dr, zebrafish (*Danio rerio*); Ol, medaka (*Oryzias latipes*); Ci, Vase tunicate (*Ciona intestinalis*); Bf, amphioxus (*Branchiostoma floridae*); Oj, feather star (*Anneissia japonica*); Lv, green sea urchin (*Lytechinus variegatus*); Sp, purple sea urchin (*Strongylocentrotus purpuratus*).

generally less derived than the earlier or later stages, indicating that the conserved phylotypic period is also less derived. We also found genes that potentially explain less derivedness, such as Hox genes. Finally, we highlight technical concerns that may influence the measured transcriptomic derivedness, such as read depth and library preparation protocols, for further improvement of our method through future studies. We anticipate that this index will serve as a quantitative guide in the search for constrained developmental phases or processes.

Keywords: derivedness, evo-devo, phenotypic evolution, phylotypic period, chordates, echinoderms

1 INTRODUCTION

Considering the fact that species of different lineages have spent exactly the same geological time since the split from their common ancestor (Baum and Smith, 2012), it can be said that they are equally evolved. However, various factors, such as different generation turnover times (Martin and Palumbi, 1993; Li et al., 1996) and population sizes (Woolfit, 2009; Lynch et al., 2016), have led to different biological times spent by them in different lineages (Bromham and Penny, 2003; Baer et al., 2007; Lanfear et al., 2010; Gaut et al., 2011). This means that the evolutionary speed of each species and lineages differs from each other, as indicated by the different genome evolutionary rates (Green et al., 2014; Urry et al., 2016). Species called living fossils, such as coelacanth (Amemiya et al., 2013) and tuatara (Miller et al., 2009), are good examples because they retain a variety of ancestral or conserved traits and have slower evolutionary rates in their genomes than their sister groups (Amemiya et al., 2013; Gemmell et al., 2020). Similarly, phenotypic changes during evolution also differ among different traits even within the same species; some traits, such as basic anatomical features, or the body plan for each animal phylum, remain strictly conserved through hundreds of millions of years (Arthur, 1997; Erwin et al., 2011), while body size or coat colors appear to change rather frequently (Stern, 2001; Hoekstra, 2006). Recent studies have demonstrated that mid-embryonic, organogenesis stages (or phylotypic period in the developmental hourglass model (Duboule, 1994)) of animals (such as vertebrates (Hazkani-Covo et al., 2005; Irie and Sehara-Fujisawa, 2007; Domazet-Lošo and Tautz, 2010; Irie and Kuratani, 2011; Wang et al., 2013; Hu et al., 2017), *Drosophila* species (Kalinka et al., 2010), nematodes (Levin et al., 2012), and molluscs (Xu et al., 2016)) are evolutionarily more conserved than their earlier or later developmental stages. This implies that conserved stages or traits are more ancestral; however, it has to be noted that conservation may not necessarily indicate that these traits are “less derived” than others, retaining more ancestral states. This is because “conservation” generally focuses on traits, genes, or genomic sequences commonly shared among the species being compared, and those that are lost or newly acquired during evolution are often excluded. In other words, the degree of changes accumulated during evolution or degree of “derivedness,” may not be effectively measured by estimations using only commonly shared features.

This causes a variety of ambiguities in understanding phenotypic evolution. For example, we still cannot determine which group of animals are more (or less) phenotypically derived on average than other groups (Irie et al., 2018). Echinoderms, for example, are often regarded as “highly derived” species (Hyman, 1955; Morris, 1999; Brusca et al., 2016), as they are a unique group of bilaterians which evolved pentaradial symmetric body plans in adults. However, the idea that echinoderms are highly derived is based on the enumeration of novel traits, such as development of body plan with pentaradial symmetry, and it still remains to be tested if echinoderms indeed experienced greater changes as a whole to their phenotypes (more derived) than their sister groups. Similarly, a detailed examination is needed to determine whether the phylotypic period is the least derived stage during development. Its conservation is often evaluated based on the expression profiles of genes that are shared in all the species (1:1 orthologs), and changes accumulated by non-shared genes are underestimated. To tackle these problems, we propose a transcriptomic “derivedness index” to address the degree of phenotypic changes observed in embryos, inherited from a common ancestor. Specifically, we compared the gene expression profiles of echinoderms (Tu et al., 2012; Li et al., 2018, 2020; Hogan et al., 2020) and chordates (Wang et al., 2013; Hu et al., 2017) to test whether echinoderms are more highly derived than chordates.

2 RESULTS

2.1 Procedure to Calculate Derivedness Index of Embryonic Transcriptomes

While our recent study indicated that echinoderms have a comparable evolutionary rate in their genomes compared with that of chordates (Li et al., 2020), it still remains to be tested if echinoderms have experienced larger modifications to their development than their sister groups. Meanwhile, it is tempting to know if developmental stages that establish pentaradial symmetry and its later stages are more highly derived than the embryos of their sister groups. To answer this question, we developed a method for quantifying the derivedness of embryos using gene expression profiles (**Figure 1A**). Specifically, we used whole embryonic developmental transcriptomes as a phenotype of embryos, and compared them between echinoderms (*Anneissia japonica*, *Apostichopus japonicus*, *Lytechinus variegatus*, *Strongylocentrotus purpuratus*) and chordates (*Mus musculus*, *Gallus gallus*, *Pelodiscus sinensis*, *Xenopus laevis*, *Danio rerio*, *Oryzias latipes*, *Ciona intestinalis*, and *Branchiostoma floridae*).

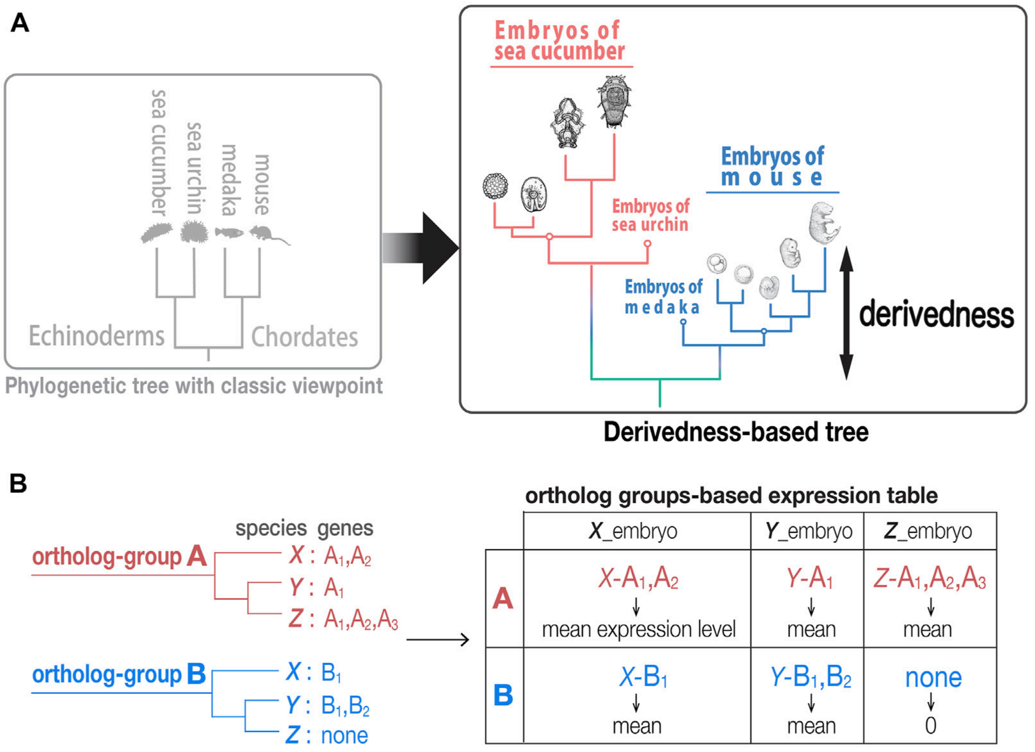

A
sea cucumber
sea urchin
medaka
mouse
Echinoderms
Chordates
Phylogenetic tree with classic viewpoint
Embryos of sea cucumber
Embryos of sea urchin
Embryos of mouse
Embryos of medaka
derivedness
Derivedness-based tree
B
species genes
ortholog-group A
X : A1,A2
Y : A1
Z : A1,A2,A3
ortholog-group B
X : B1
Y : B1,B2
Z : none
ortholog groups-based expression table
X_embryo
Y_embryo
Z_embryo
A
X-A1,A2 → mean expression level
Y-A1 → mean
Z-A1,A2,A3 → mean
B
X-B1 → mean
Y-B1,B2 → mean
none → 0

FIGURE 1 | Tree based on phenotypic derivedness of embryos. **(A)** Schematic illustration of a phylogenetic tree drawn by classic viewpoint (left) and that by transcriptomic derivedness (right). Phenotypic derivedness aims to introduce quantitative evaluation of how derived each sample or embryo became from their common ancestor. **(B)** In contrast to evaluating conservation with commonly shared genes, such as 1:1 orthologs, our method takes advantage of ortholog group-based expression table, which considers the expression of paralogs and potential lost genes.

In contrast to approaches that compare the expression of only conserved 1:1 orthologs, our approach encompasses expression of paralogs and potentially lost genes to cover as many evolutionary changes as possible. As indicated in a previous study, 1: 1 ortholog-based comparisons encounter a major barrier when multiple species are compared (Hu et al., 2017). In brief, the number of 1:1 orthologs that can be identified comes up to a very small ratio of genes in the entire genome (e.g., only 1,704 1:1 orthologs could be identified in their analysis with eight chordate species (Hu et al., 2017)). This situation is intensified in our analysis when both chordates and echinoderms are involved (13 species in total), where only 271 1:1 orthologs could be identified. These not only account for ~1% of all genes in a typical vertebrate genome (~20,000 genes) but also overlook the changes (such as those by gene duplication and/or gene loss) that occur during evolution, which leads to possible underestimation of how derived each embryo is from their common ancestor.

To identify genes in orthologous groups of distantly related species, we first compared protein-coding genes and identified 22,699 ortholog-groups using the PorthoMCL software for eight chordates, four echinoderms, and one outgroup species (Li et al., 2003; Tabari and Su, 2017) (**Supplementary Figure S1**). We then calculated the normalized expression (see also **Table 1** in Methods for the normalization methods tested) for each ortholog-group by 1) taking the mean expression of paralogs and 2) giving "zero" expression value to potentially lost genes (taking sum-expression of paralogs also provided similar results, as in a previous study (Hu et al., 2017); see also **Figure 1B** and **Supplementary Figure S3A**). However, species-specific genes were not included, as the analyses including these genes did not meet the criteria we utilized (described in detail below).

Finally, transcriptomic distance was calculated with various combinations of normalization, distance, and tree inference methods and scanned for a suitable combination (**Figure 2A**) based on the following three criteria (**Figure 2B**): 1) developmental stages cluster by species (Kalinka et al., 2010; Hu et al., 2017); 2) tree topology is consistent with known phylogeny inferred from genomic sequences (suggested by previous studies (Kalinka et al., 2010; Hu et al., 2017)); and 3) within-species transcriptomic distances show gradual changes along development (high smoothness in the distance image).

Criterion 1, a topology of developmental stages clustered by species, can be expected for the species covered in this study. For example, even with six closely related *Drosophila* species that split less than 40 million years ago and share very similar morphological features, their whole embryonic transcriptomes still cluster by species rather than by stages (Kalinka et al., 2010). This is presumably because despite sharing many conserved developmental features, larger differences including those with respect to species-specific characteristics exist, such as ovariole

TABLE 1 | Derivedness tree construction method combinations for scanning.

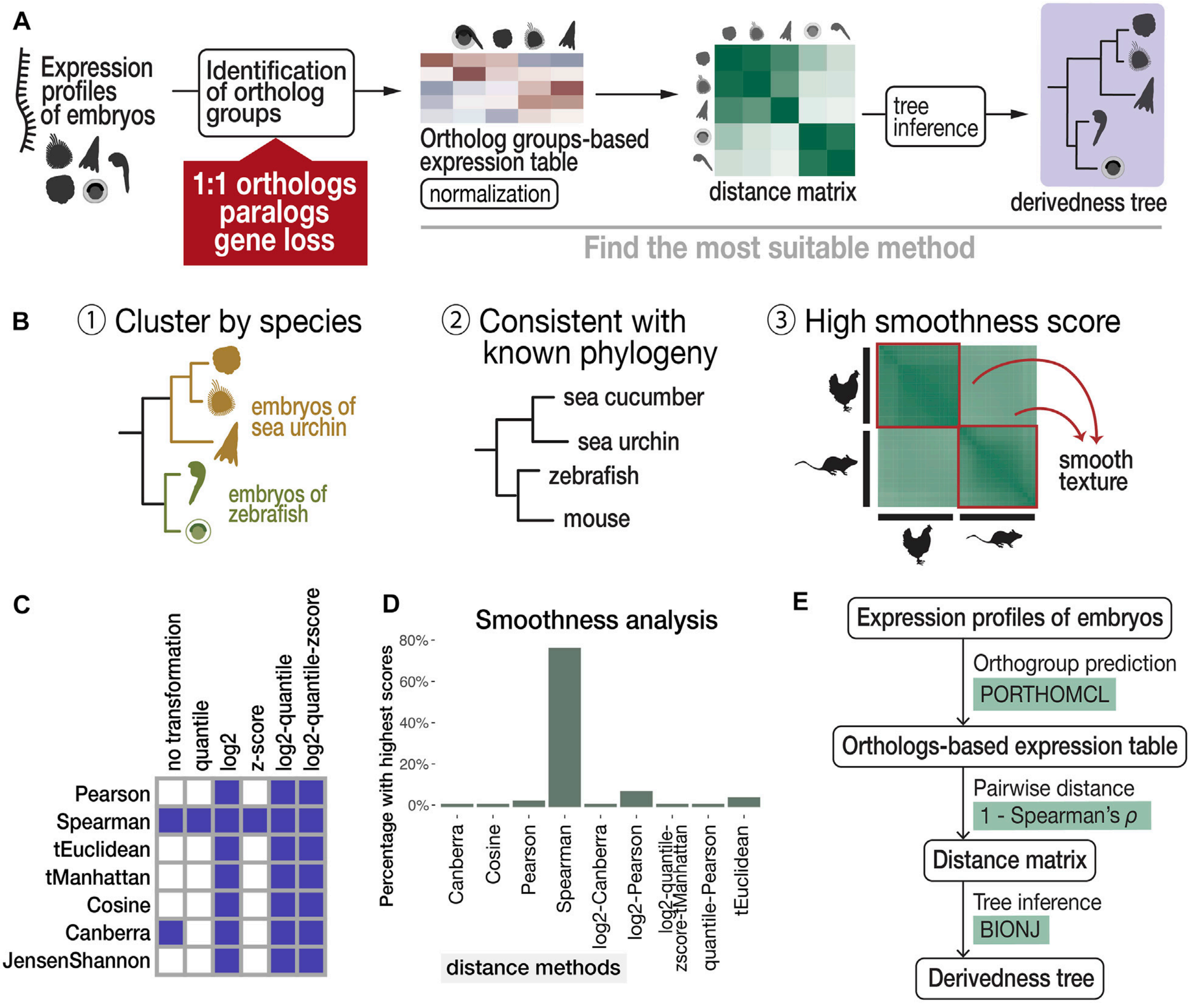
Procedure	Methods to test
Normalization	$\log_2$ (TPM+1) $\log_{10}$ (TPM+1) quantile normalization ascending rank descending rank z-score $\log_2$ (TPM+1) → quantile $\log_2$ (TPM+1) → z-score $\log_2$ (TPM+1) → quantile → z-score
Distance methods	1—Pearson's correlation coefficient r 1—Spearman's correlation coefficient ρ tEuclidean tManhattan Cosine Canberra Jensen-Shannon
Tree inference	NJ (neighbor-joining) (Saitou and Nei, 1987) BIONJ (Gascuel, 1997) FastME, balanced (Lefort et al., 2015) FastME, OLS (Lefort et al., 2015) Fitch-Margoliash (PHYLIP) (Fitch and Margoliash, 1967; Revell and Chamberlain, 2014)

number and genome size (Markow and O'Grady, 2007), and this may have clustered samples by species. Development-related genes account for only 10–25% of the whole genome of chordate species (Hu et al., 2017). Therefore, in this study, we selected representative species that should be distantly related enough to avoid the tree topology of clustering by stage. Criterion 2 was based on the assumption that the derivedness of the tree based on overall transcriptomic profiles does not differ from the known phylogenetic topology deduced from the genome, and this is supported by a study that analyzed more closely related *Drosophila* species (Kalinka et al., 2010). For criterion 3, we deployed image texture analysis to score the degree of smoothness of distance matrix images generated by different distance methods, as transcriptomic changes are expected to be continuous along developmental stages.

The results of our scanning show that both rank and logarithmic transformations of the expression data met criterion 1, namely, clustering samples by species (shown in **Figure 2C** and in the visualized expression images in **Supplementary Figure S2**). In addition, we found that calculating pairwise distances of gene expression profiles using "1—Spearman's correlation coefficient ρ" (abbreviated as *Spearman distance* below) meets all three criteria and is the most suitable method among the tested distance methods (**Figures 2C**, **2D**, **3A**). Finally, we tested several commonly used tree inference methods, where BIONJ outperformed the other tested methods with its speed and algorithmic design to cluster samples with the lowest distances together (Saitou and Nei, 1987; Gascuel, 1997), although most methods generated similar results except for the Fitch-Margoliash method (**Supplementary Figure S3B**). A simplified scheme of the most suitable combination based on our criteria is summarized in **Figure 2E**. As mentioned above, our analyses did not cover the expression of species-specific genes. Intuitively, inclusion of species-specific genes may allow us to more comprehensively cover and evaluate the changes made during evolution; however, we excluded these genes as the analyses including species-specific genes did not meet criterion 2 (consistent with known phylogeny). *X. laevis* became an outgroup of other vertebrates in this tree (**Supplementary Figure S4**). This may be due to the low accuracy of gene prediction, which led to an excessive number of species-specific genes which biases the distance calculation (**Supplementary Figure S1D**). Notably, 7,879 (40%) of these non-paralogous, specific-specific genes in *X. laevis* are lowly expressed (max TPM <1) throughout the developmental stages examined, suggesting that many of these genes might be inaccurately annotated.

2.2 Penta-Radial Phase of Echinoderms Appear to Be Highly Derived, but Their Overall Developmental Systems May Not Be Much More Derived Than Those of Chordates

We defined the derivedness index of each embryo as the total branch length from the putative common ancestral node on the inferred tree (**Figure 3A**), which is the common ancestor of chordates and echinoderms. As has been assumed, some of the echinoderm species, especially embryos of two sea urchins, showed much higher derivedness indices than those of invertebrate chordates. Notably, the tendency was more obvious for the developmental stages with pentaradial symmetry (shaded stages in **Figure 3A**). Meanwhile, the other echinoderms, such as the feather star (an early-diverged echinoderm species (Cannon et al., 2014; Li et al., 2020)) and the sea cucumber showed rather less-derived indices when compared to those of vertebrates (**Figure 3A**). These tendencies were corroborated by summarizing the range of the derivedness indices of all embryos for each species [**Figure 3B**; changes in the indices among different species were statistically significant (p < 2e-16 by Kruskal-Wallis test)], and that for all chordates and echinoderm species. (**Figure 3C**). Nevertheless, when we compared the derivedness index of the penta-radial phase of echinoderm development (including metamorphosis, juvenile, and adult stages) against the bilateral phase (early stages to larval stages, before penta-radial structures start to appear) in each species, the penta-radial phase was indeed more derived in the feather star and the two sea urchin species (**Figure 3D**, statistically significant, tested with the Wilcoxon signed-rank test). Considering that the two sea urchin species, which split around 180 million years ago (Li et al., 2020), share similar developmental characteristics, the reason why *L. variegatus* showed a relatively higher derivedness index than *S. purpuratus* remains to be clarified (*L. varietagus* became even more highly derived when species-specific genes were considered; **Supplementary Figure S4**). However, this could partially be due to *L. variegatus* having a faster genomic evolutionary rate than *S. purpuratus* (Li et al., 2020; see also **Supplementary Figure S19**). These results suggest that although echinoderms may not have experienced larger modifications to their molecular developmental systems than chordates since their split from the deuterostome common ancestor, sea urchins appear to


A
Expression profiles of embryos
Identification of ortholog groups
1:1 orthologs
paralogs
gene loss
Ortholog groups-based expression table
normalization
distance matrix
tree inference
derivedness tree
Find the most suitable method
B
① Cluster by species
embryos of sea urchin
embryos of zebrafish
② Consistent with known phylogeny
sea cucumber
sea urchin
zebrafish
mouse
③ High smoothness score
smooth texture
C
D
Smoothness analysis
Percentage with highest scores
80%
60%
40%
20%
0%
Canberra
Cosine
Pearson
Spearman
log2-Canberra
log2-Pearson
log2-quantile-zscore-tManhattan
quantile-Pearson
tEuclidean
distance methods
E
Expression profiles of embryos
Orthogroup prediction
PORTHOMCL
Orthologs-based expression table
Pairwise distance
1 - Spearman's ρ
Distance matrix
Tree inference
BIONJ
Derivedness tree

	no transformation	quantile	log2	z-score	log2-quantile	log2-quantile-zscore
Pearson	☐	☐	■	☐	■	■
Spearman	■	■	■	■	■	■
tEuclidean	☐	☐	■	☐	■	■
tManhattan	☐	☐	■	☐	■	■
Cosine	☐	☐	■	☐	■	■
Canberra	■	☐	■	☐	■	■
JensenShannon	☐	☐	■	☐	■	■

FIGURE 2 | Proposed workflow to estimate phenotypic derivedness of embryos using gene expression profiles and evaluation of methods. **(A)** Outline of measuring transcriptomic derivedness. Whole embryonic transcriptomes from different species were compared using the expression levels of ortholog groups, which consider not only 1:1 orthologs but also paralogs and genes that are presumably lost in specific lineages. Derivedness index of each embryo was defined as the total branch length from the putative common ancestral node on the inferred tree. **(B)** Criteria for selecting the suitable method for quantifying derivedness of embryos. These include: [1] clusters samples by species; [2] topology of derivedness tree consistent with known phylogeny estimated from genomic sequences (Kalinka et al., 2010; Hu et al., 2017), with support from biological replicates; [3] Transcriptomic similarities show gradual change along developmental stages. **(C)** Rank (involved in calculating Spearman distance) and logarithmic normalizations of expression data tend to meet the criteria of clustering stages by species. Shaded boxes represent normalization and distance methods showing this topology in the inferred tree. **(D)** Spearman distance scores the highest in smoothness analysis. (E) Summary of the most suitable method among the combinations tested.

have accumulated more changes in their developmental systems, especially in the later embryonic stages when pentaradial structures become evident. However, this may not be solely due to gene expression from pentaradial structures or their source structures, as the sample "Lv_8wpfRudiment" (the forming rudiment in 8-weeks-post-fertilization embryo of *L. variegatus*) is almost as highly derived as "Lv_8wpfLarva" (the remaining larval body with the rudiment removed), which does not contain pentaradial structures. The differences in transcriptomic derivedness indices of the pre-metamorphosis and the pentaradial phases could at least partly be attributed to how genes are expressed at different levels in the two phases (rather than the two developmental phases deploying different sets of genes) (**Supplementary Figure S17**). Further studies delineating the molecular mechanisms of the metamorphosis process may help explain why stages with pentaradial structures tend to show highly derived indices.

2.3 "Conserved" Phylotypic Periods Are the Least Derived in Vertebrates

Consistent with previous studies that supported the developmental hourglass model with 1:1 orthologs (Irie and Kuratani, 2011; Levin et al., 2012; Wang et al., 2013; Uesaka

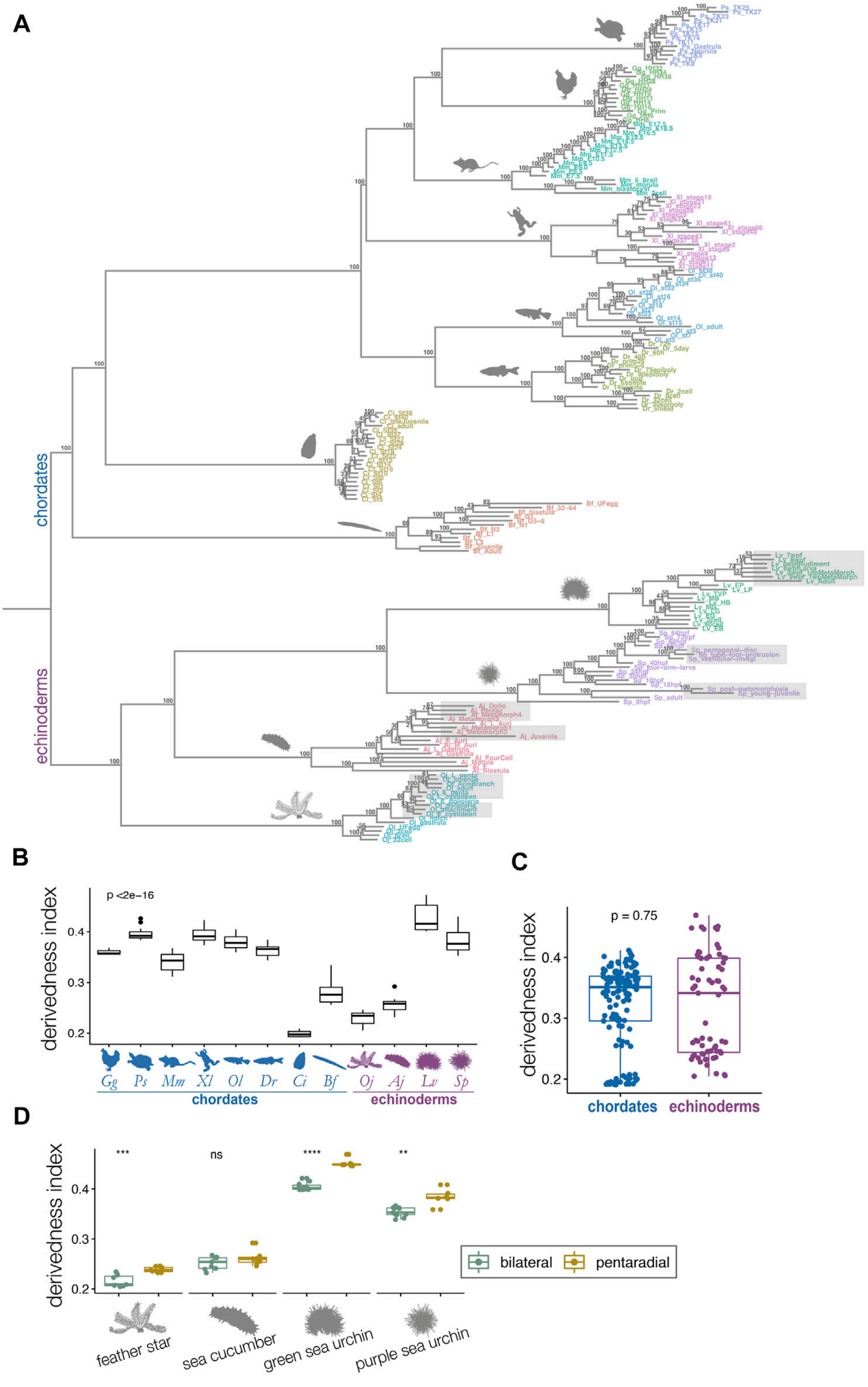


FIGURE 3 | Derivedness tree of developmental stages of echinoderms and chordates based on transcriptomic profiles. **(A)** Tree showing transcriptomic derivedness of chordate and echinoderm embryos. Embryos of each species are denoted by the same color. Support values are consensus from 100 random biological replicates-included (BRI) trees (see also Methods). Shaded stages in echinoderm species are developmental phases with penta-radial structures. **(B)** The range of derivedness indices of embryos for each species. Vertebrate species and the two sea urchin species have fairly high derivedness indices whereas *C. intestinalis* has the lowest (blue: chordate species; purple: echinoderm species; Kruskal-Wallis $p < 2e-16$) **(C)** The range of derivedness indices of echinoderm and chordate embryos. This shows that echinoderms do not appear to have much higher derivedness index than chordates. (Mann-Whitney $p = 0.75$) **(D)** The range of derivedness indices of the

(Continued)

FIGURE 3 | bilateral (green) and the pentaradial (yellow) phases of echinoderm development. The penta-radial phase is more derived in feather star (*Oj*), green sea urchin (*Lv*), and purple sea urchin (*Sp*). (Mann-Whitney *U* test, **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$; ns: $p > 0.05$). Species abbreviations: *Gg*, chicken; *Ps*, soft-shelled turtle; *Mm*, mouse; *Xl*, clawed frog; *Ol*, medaka; *Dr*, zebrafish; *Ci*, ascidian; *Bf*, amphioxus; *Oj*, feather star; *Aj*, sea cucumber; *Lv*, green sea urchin; *Sp*, purple sea urchin.

et al., 2019), our results show a similar trend in vertebrates, with conserved mid-embryogenesis (especially organogenesis stages) being the least derived within each species from the common ancestor of chordates (**Figures 4A–F**). Meanwhile, in some species, the identified least derived stages showed a slight shift from the previously reported conserved stages; for instance, while Prim-5-6 of the zebrafish (*D. rerio*) was previously suggested to be the most conserved stage (or vertebrate-phylotypic period), but the least derived stage was at a slightly earlier, 14-somites stage, and the Prim-5-6 stage was the second least derived embryo (see also **Supplementary Figure S7**). Larger differences were observed in *Ciona* and amphioxus. While the most conserved stages, identified by 1:1 orthologs, in *Ciona* were around st.24–29, the least derived stages were st.1–10. In amphioxus, while the mid-neurula stage (Hu et al., 2017; Marlétaz et al., 2018) was the most conserved, stage L2 (open-mouth larva) was found to be the least derived. These results imply that "conserved" phylotypic period are in general less derived than the earlier/later stages, but these may show larger differences especially when more number and/or distantly related species are being compared.

Nevertheless, among the species analyzed, for the range of derivedness indices and position of the common ancestral node of all the embryos of each species, the Chinese soft-shell turtle tends to show more derived developmental systems than the other vertebrates. This is consistent with the view that turtles possess many highly derived morphological diapsid features (Nagashima et al., 2009; Gilbert and Corfe, 2013). However, genomic analyses indicated that turtles have a slow evolutionary rate (Bradley Shaffer et al., 2013; Green et al., 2014), and further studies are required to explain this discrepancy. Meanwhile, given that the fossil records of turtles remained morphologically conserved since their appearance (Li et al., 2008; Benton, 2014), it is possible that the turtle embryos are transcriptomically more derived (expressing orthologous genes in a different way) than the other vertebrate species while maintaining a slow genomic evolutionary rate. In contrast, our results indicate that the mouse appears to possess the least derived molecular developmental system among the vertebrates compared (**Figures 3B**, **6B-iii**). Another unexpected result was that the overall stages of *Ciona* showed the least derived indices when compared with those of the embryos of other chordates, including the early diverged amphioxus. This is in contrast to both morphological and genomic studies, which suggested that tunicates could be the most derived species in chordates (Holland, 2015, 2016). Moreover, a similar tendency was also corroborated by the tree that covered species-specific genes (**Supplementary Figure S4**) and the tree drawn by 1:1 orthologs (**Supplementary Figure S5**). The exact reason for this is unclear; however, a potential reason would be due to the vast number of potentially lost genes in *Ciona*, as it may bias the distances against different species by giving higher transcriptomic similarity to the rest of the species. Given that our method correctly captured transcriptomic derivedness, the results imply that their orthologous genes used during embryonic development remain rather ancestral, despite their highly divergent genomic sequences (Berná and Alvarez-Valin, 2014) after the split from the Olfactores common ancestor.

2.4 Characterization of Least Derived Stages by Extracting Highly Expressed Ortholog-Groups

While there were a few reports that hint how the body plan-establishing phase became conserved (Bogdanović et al., 2016; Hu et al., 2017; Zalts and Yanai, 2017; Uchida et al., 2018), we further sought for potential hints toward the mid-embryonic conservation. Specifically, we asked which ortholog-groups could potentially characterize this conserved phase, given that the least derived stage only differed slightly from the reported conserved phase in vertebrates. For this purpose, we sought ortholog-groups that were highly/lowly expressed during the stages with less derivedness, by calculating the correlation (Spearman's correlation coefficient) between the expression of each ortholog-group and derivedness indices during development. For example, ortholog-groups with strong negative correlations across species (dark blue in **Figure 5A**, abbreviated as "negative DCOs," derivedness-correlative ortholog-groups) tended to show higher expression around the less derived developmental phase and lower expressions in more derived stages. We especially looked for negative DCOs across the six vertebrate species, as they may represent transcriptomic features of the phylotypic period and could potentially provide a hint about the mechanism of its conservation.

To avoid unexpected bias, we excluded samples obtained by Quartz-Seq (2-cells to blastocyst in mouse and all stages in sea cucumber) for this analysis, as these were obtained by different RNA-seq protocols than the others (discussed in detail in the next section). Among the 22,699 ortholog-groups analyzed, we focused on 7,775 ortholog-groups that had at least one gene counterpart in each vertebrate species (**Figure 5B**). Among these, we found 695 ortholog-groups showing negative correlation in all six vertebrate species concomitantly (category 6 in **Figure 5B**; **Supplementary Table S19**), including 18 HOX ortholog-groups, 201 development-related ortholog-groups, and 161 ortholog-groups involved in signaling transduction (**Figure 5C**, **Supplementary Figures S8–10**; the ratio of development-related ortholog-groups is significantly higher in category 6 than in any other category; Fisher's exact test), consistent with the actively proceeding organogenesis in these stages. Along with whole embryonic expression, some of aforementioned genes are also known to show conserved spatial expression patterns (Thisse et al., 2004; Darnell et al., 2007; Yokoyama et al., 2009; Bowes et al., 2010;

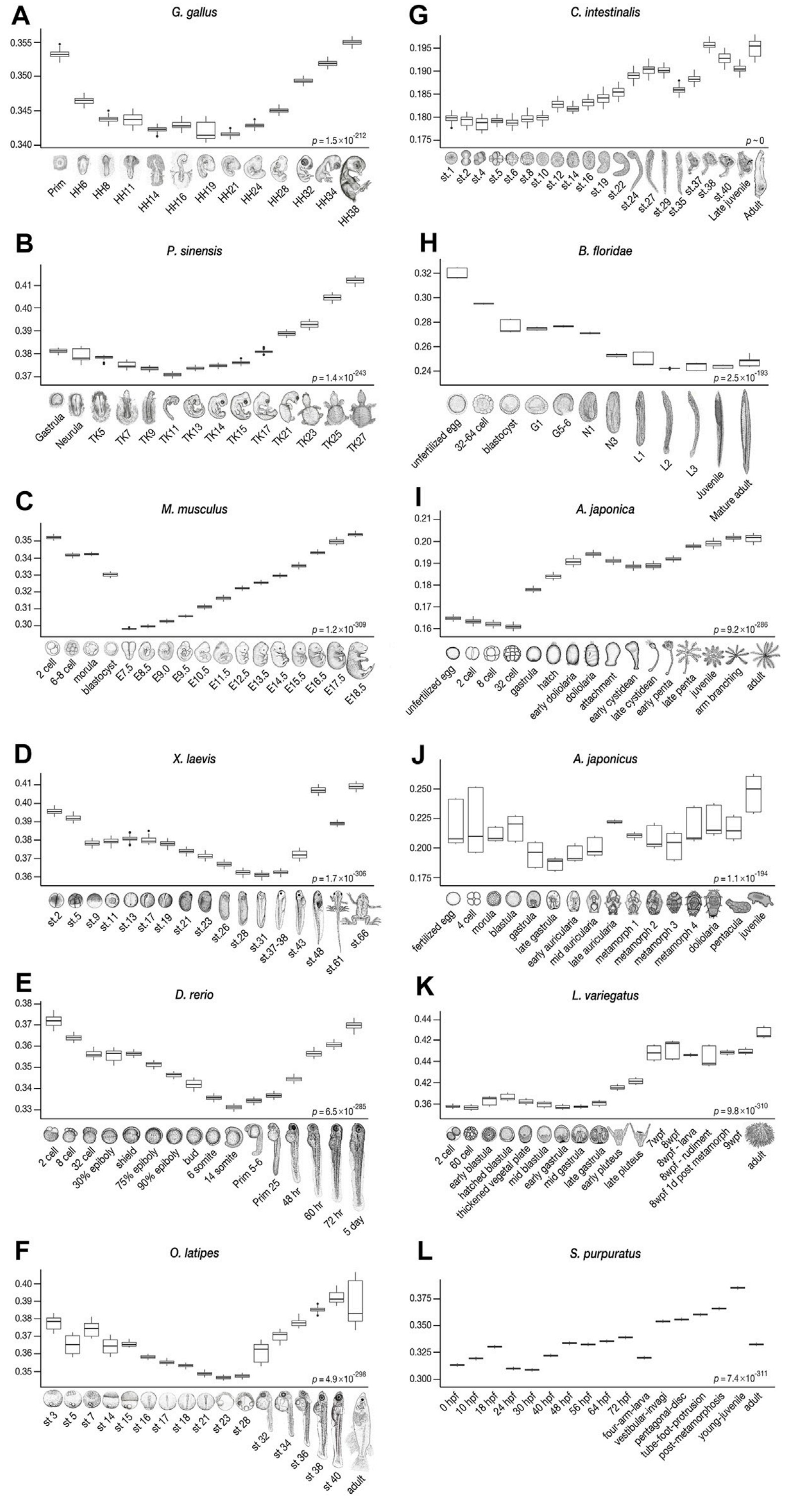


FIGURE 4 | Derivedness index of embryos of chordates and echinoderms estimated from their respective common ancestors. The range of derivedness indices of each embryo, from the common ancestor of chordates and echinoderms, respectively, in 100 biological replicates-included (BRI) trees. **(A–F)** The least derived stages in vertebrate species (*Gg*, *Ps*, *Mm*, *Xl*, *Dr*, and *Ol*) are mid-embryonic and organogenesis stages (*Gg*: HH14-19, *Ps*: TK11, *Mm*: E7.5*, *Xl*: stage 31, *Dr*: 14-somites, and *Ol*: stage 23). *: E9.5, in mouse when the Quartz-Seq samples were removed (see **Figures 6B–6ii**). **(G–H)** The tunicate *C. intestinalis* shows relatively lower derivedness indices in stage 1 to stage 10 embryos, and the least derived stage in the amphioxus *B. floridae* is around the L2 (open-mouth larva) stage. **(I–L)** In echinoderm species, the least derived stage is around the gastrula in sea cucumber and sea urchins, whereas the 32-cell stage is the least derived in feather star. (Differences in derivedness index for each developmental stage are statistically significant; Friedman test).

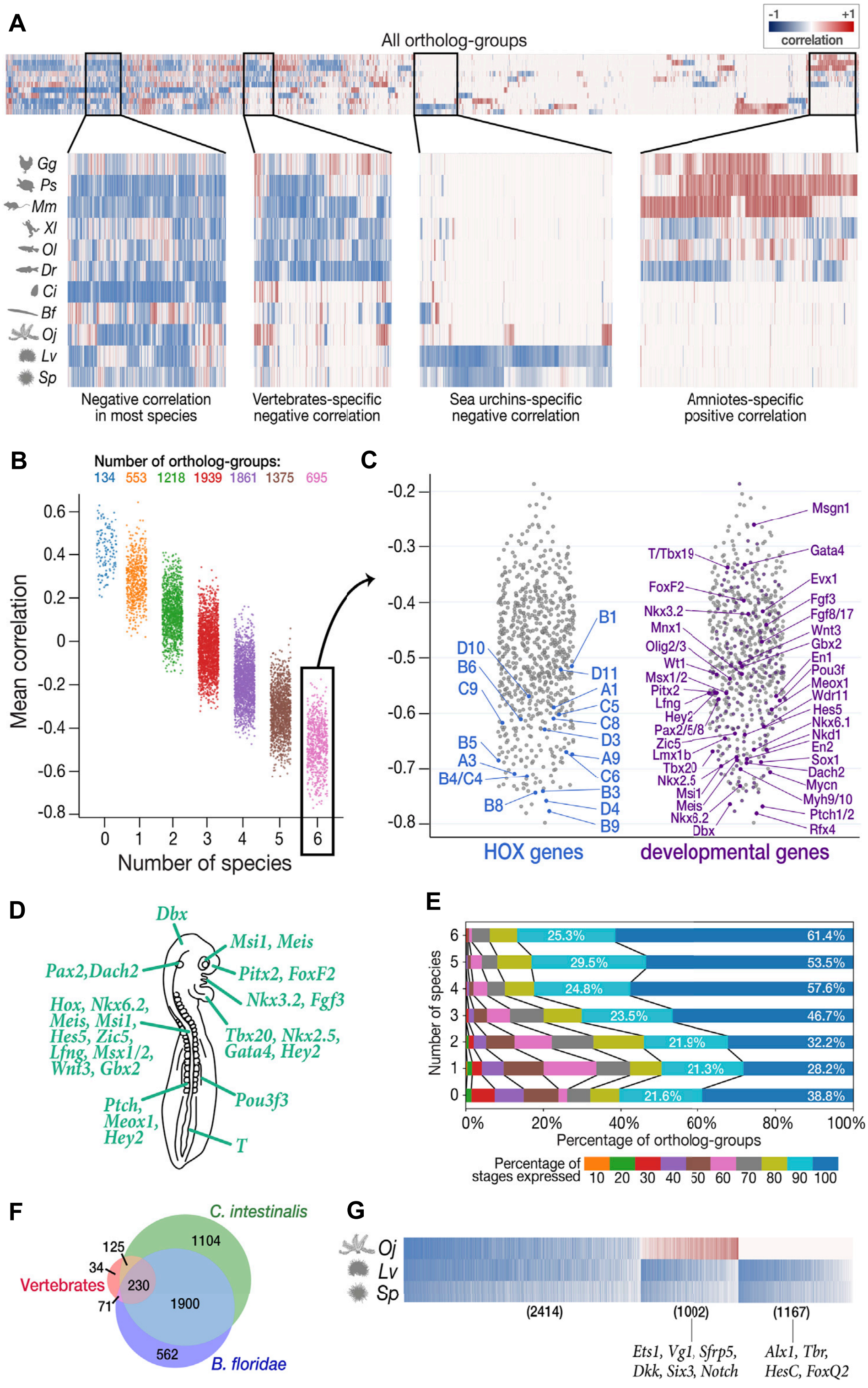


FIGURE 5 | Characterization of least derived stages by ortholog-groups. **(A)** Heatmap showing Spearman's correlation coefficient between expression levels and derivedness indices for each ortholog-group for each species (blue: negative correlation; red: positive correlation). Selected areas in the heatmap are zoomed in to show clusters of ortholog-groups with (from left to right) negative correlation in most species; negative correlation in most vertebrate species; negative correlation only in the two sea urchins; positive correlation in amniotes (mouse, chicken, and turtle). **(B)** Ortholog-groups showing negative correlation (negative DCOs) in vertebrates. 7,775 ortholog-groups were analyzed for vertebrates, and were further classified into seven categories based on the number of species in which they show negative correlation. 695 ortholog-groups in category 6 (showing negative correlations in all six vertebrate species analyzed) were further analyzed in **(C–E)**. **(C)** Category 6 *(Continued)*

FIGURE 5 | negative DCOs, including 18 HOX ortholog-groups (left) and 201 development-related ortholog-groups (right). **(D)** The putative phenotype of the least derived/conserved mid-embryonic stage of vertebrates. Genes known to express in homologous anatomical structures during this developmental phase in mouse, chicken, frog and zebrafish are highlighted. **(E)** DCOs with negative correlations in more species tended to show higher degree of temporal pleiotropy, estimated by the ratios of stages detected (TPM ≥ 1). **(F)** The number of DCOs with negative correlations in 6 vertebrate species (red), *C. intestinalis* (tunicate; green), and *B. floridae* (amphioxus; blue). 230 ortholog-groups showed negative correlations in all three groups, suggesting that they might be involved in ancestral functions retained in chordates. **(G)** 2,414 negative DCOs were detected across echinoderm species (left); however, the functions of most of these genes remain unknown. Well-studied gastrulation-related genes, such as *Ets1* and *HesC*, were among the groups showing positive correlation in feather star (middle) and the sea urchin-specific gene group (right), respectively.

Richardson et al., 2014; Smith et al., 2018) (**Figure 5C**). Of note, these negative DCO genes that show spatial conservation contain many genes that are reported to be involved in neural patterning. For example, *Nkx6.1*, *Pax2*, *Lmx1b*, *Wt1*, *Evx1/2*, *En1*, *Mnx1*, *Sox1*, and *Olig2/3* are involved in the dorsal-ventral patterning of neurons in the spinal cord (Catela et al., 2014). Additional examples are the two ortholog-groups with the strongest mean negative correlation. *Patched1/2* and *Rfx4*, which are receptors of sonic hedgehog (Shh) and a downstream target of Shh signaling, respectively, whose mutants show severe defects in the neural tube (Goodrich et al., 1999; Ashique et al., 2009; Murdoch and Copp, 2010; Sedykh et al., 2018) (see also **Supplementary Figure S11**; Shh also showed negative correlation across species). Genes involved in other conserved structures of the phylotypic embryo were also identified as negative DCOs (**Figures 5C,D**), such as those involved in somitogenesis (*Patched1/2*, *Nkd2*), heart (*Nkx2.5*, *Tbx20*, *Gata4*), and mesonephros (*Pou3f3*). Besides, this list of negative DCOs also included genes that were not previously considered to be directly related to embryonic development, such as *Fidgetin* (the ortholog-group showing the strongest mean negative correlation of −0.798 among all ortholog-groups analyzed). *Fidgetin* has been recently suggested to regulate the cytoskeleton in the spinal cord and somites in the mouse embryo (Leo et al., 2015). More genes that are involved in organogenesis could potentially be identified from these DCOs.

In addition, consistent with a recent report suggesting that developmental stages with more pleiotropic (repetitively used) genes tend to be evolutionarily conserved, ortholog-groups showing negative correlation in more species tended to be expressed in more developmental stages (temporal pleiotropy, **Figure 5E**), suggesting that negative DCOs are potentially useful for characterizing the least derived stages. However, we note that this method may overlook ortholog-groups that have significant contribution to the phenotype of the least derived stages but show consistently high expression throughout development because they may not show a strong negative correlation with the derivedness index.

Finally, we extended the analysis to the entire chordate and echinoderm clades, which have been difficult to perform with 1:1 ortholog-based or conservation-oriented methods (as only ~1% protein-coding genes of the entire genome could be compared). Among the 4,026 chordate ortholog-groups, 230 were identified as negative DCOs in all eight chordate species analyzed (**Figure 5F**, **Supplementary Table S20**). As expected, these contained a variety of genes that are known to be important for chordate embryogenesis, such as *Tbx20* (Belgacem et al., 2011), which showed a strong average negative correlation. This implies that negative DCOs might provide hints for the identification of genes retaining conserved functions shared in chordates in future studies, however, it also has to be noted that sequence-based similarity does not always indicate their similar functions, and roles of negative DCOs have to be analyzed carefully in the future studies. Similarly, in echinoderms, we found 2,414 DCOs showing negative correlation across all three echinoderm species (left-most in **Figure 5G**, **Supplementary Table S21**), which contained genes mainly expressed in gastrula in sea urchins and 2-cell to 32-cell stage in feature star (the least derived stages shown in **Figure 4I–L**). Since the functions of genes in echinoderm species remain largely unexplored, we could not deduce much information from the negative DCOs in echinoderms; however, some implications were obtained for smaller groups in echinoderms, namely, sea urchins (**Figure 5G**). In contrast, genes involved in the establishment and growth of the pentaradial phase were included in the cluster showing positive correlation in all three echinoderm species since this developmental phase was shown to be more derived.

2.5 Technical Concerns in Cross-Species Transcriptomic Analyses

While many of our results correspond with previous EvoDevo studies, it has to be noted that there are several technical issues to overcome. First, the choice of outgroup species may affect the topology of the tree, as is the case in most evolutionary studies. In contrast to **Figure 3A**, the tree with *Drosophila melanogaster* as the outgroup (**Supplementary Figure S13**) showed a tree topology that deviated more from the known phylogeny. This could potentially be because fewer genes in *D. melanogaster,* (which is often considered a highly derived species in arthropods (Andrioli, 2012), could be identified as orthologous genes of deuterostomes. Second, differences in read depth may cause biases in expression quantification, as samples with deeper sequencing are expected to detect more lowly expressed genes than shallower samples. To test how these may affect the derivedness index, additional trees were plotted from expression data with depth adjusted uniformly to 10 million (randomly picked up 10 million mapped reads, 10 M; which corresponds to the depth of the shallowest sample), gradual down-sampling of read depths, or adjusted proportionally to the exome size of each species. In the tree with the 10 M expression data, the range of derivedness indices for six species, notably *Ciona*, increased significantly (**Figures 6A–i**, see also **Supplementary Figures S14**, **S16**), while the tree with exome sizes-adjusted depths showed significant changes in fewer species (**Supplementary Figure S15**). Notably, the topologies of both trees were inconsistent with known phylogeny. This suggests

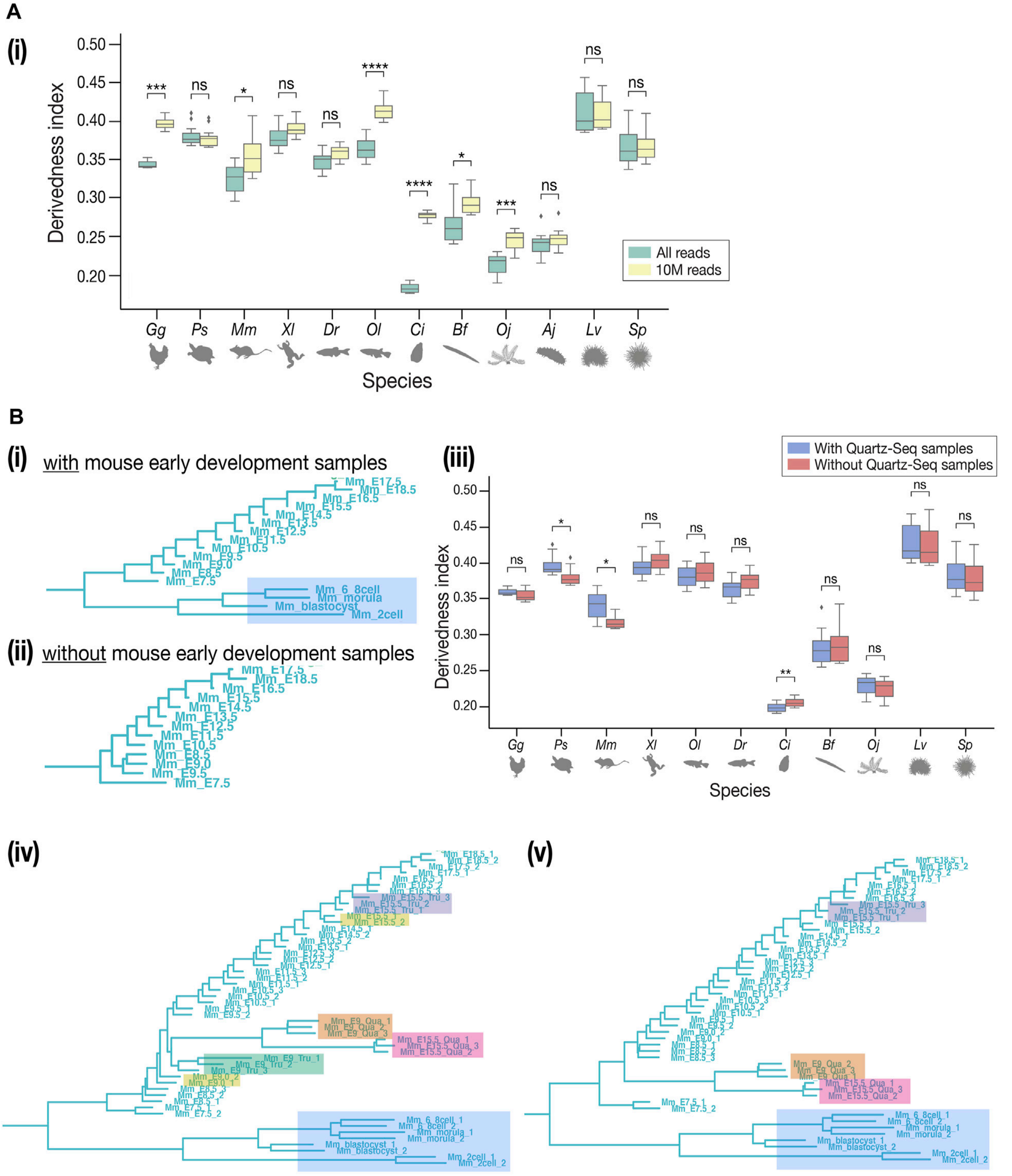


FIGURE 6 | Technical concerns in derivedness tree inference. **(A)** Read depth control. (i) In the tree inferred from 10 M-controlled expression data, the range of derivedness indices significantly increased in some species (*Gg*, *Mm*, *Ol*, *Ci*, *Bf*, and *Oj*). **(B)** Comparison of TruSeq and Quartz-Seq datasets. **(i)** The samples of early development in mouse (shaded in blue) and all samples of sea cucumber (not shown) were collected by Quartz-seq. **(ii)** Removing the Quartz-seq samples changed the tree topology inside the mouse clade. The least derived stage became E9.5. **(iii)** After removing the Quartz-seq samples, the range of derivedness indices for most species were not greatly affected, not influencing the conclusions we drew in the previous sessions. **(iv–v)** Tree with all original samples and the new samples to compare TruSeq & Quartz-seq datasets. Unexpectedly, the samples tend to cluster by protocols rather than by stages or biological replicates. This tendency is stronger in (v) where the E9.0 TruSeq samples were omitted, while the E15.5 Quartz-seq and Tru-Seq samples still did not cluster together. Samples from the same starting total RNA (*Continued*)

FIGURE 6 | but by different protocols (such as E15.5_Tru_1 and E15.5_Qua_1) did not cluster together either. (Shaded in orange: E9.0 by Quartz-seq; green: E9.0 by TruSeq; pink: E15.5 by Quartz-seq; purple: E15.5 by TruSeq; yellow: E9.0 and E15.5 by TruSeq from the published dataset). For **(A-i)**, **(A-ii)**, and **(B-iii)**, Mann–Whitney–Wilcoxon test was performed (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$; ns: $p > 0.05$).

that although the derivedness index tends to be influenced by read depth, the influence may be dependent on the species or the samples; thus, more comprehensive studies are needed to determine which read depth normalization method would be more suitable for measuring transcriptomic derivedness. Finally, as we included samples acquired by different library preparation protocols where the Quartz-Seq protocol involved an additional whole-transcript amplification step (Sasagawa et al., 2013), we asked whether these samples were comparable to each other. To answer this, we collected new RNA-seq data from E9.0 and E15.5 mouse embryos by performing both TruSeq and Quartz-Seq from the same starting total RNA sample. A tree with all samples, including biological replicates, was plotted. Unexpectedly, samples tended to cluster by protocols rather than by stages or biological replicates (**Figures 6B-iv,v**), suggesting that samples obtained by these two protocols may not be completely comparable. To avoid this bias, we excluded samples acquired by Quartz-Seq for the analysis of derivedness-correlative ortholog-groups.

3 DISCUSSION

The quantitative concept of the degree of phenotypic evolution, or phenotypic derivedness, originates from the early history of evolutionary thoughts when taxonomists attempted to compare degree of evolution of traits (Mayr, 1982), but further development of methods under modern evolutionary theory was mostly not pursued. This could partly be owing to multiple factors or understandings, such as *scala naturae* of the pre-Darwinian era (Diogo et al., 2015) and Haeckel's recapitulation theory [reviewed in (Kuratani et al., 2021)]. We contend that rather than simply abrogating this idea or mixing it with the concept of conservation, the quantitative degree of phenotypic evolution, or derivedness, may help us more deeply understand phenotypic evolution. As an application, we developed a transcriptomic derivedness index that considers not only the expression levels of strictly conserved 1:1 orthologs, but also those of paralogs and potentially lost genes. This contrasts with previous approaches, which only focused on strictly conserved 1:1 orthologs, whereas our method covers most of the genes in the genome of each species. While species-specific genes could not be included since the analysis including these genes violated one of the criteria set for this study, our method is still advantageous than the 1:1 orthologs-based method. This is because the differences in the genes being covered for each species become extremely small in the previous method, especially when a large number of species are being compared. Our scanning demonstrated that calculating evolutionary distance between embryonic transcriptomes using Spearman distance meets the criteria, including phylogenetic topology estimated by transcriptomic information recapitulates that estimated by genomic information, and the estimated tree topology show "developmental stages cluster by species".

Using this transcriptomic "derivedness index," we quantified the derivedness of whole embryonic transcriptomes by utilizing gene expression profiles of echinoderm and chordate embryos, and tested whether echinoderms are highly derived species (Hyman, 1955; Morris, 1999; Brusca et al., 2016). Unexpectedly, the tree (**Figure 3**) suggested that developmental systems of echinoderms might not have experienced larger modifications than those of chordates since the split from their common ancestor, the penta-radial phase of echinoderm species tends to be highly derived, as had been assumed. Meanwhile, in the vertebrate clade, we found that conserved mid-embryonic stages (the phylotypic period) in vertebrates tended to be less derived as well (**Figure 4**). In contrast to the situation in vertebrates, larger differences between the least derived stage and the most "conserved" stage were observed in the tunicate *C. intestinalis* and the amphioxus *B. floridae*; however, the reason is unclear. The least derived stages in *C. intestinalis* (cleavage) do not seem to span developmental phases responsible for the conserved anatomical structures of chordates [including notochord, pharyngeal gill slits, dorsal nerve cord, and segmental muscles (Benton, 2014; Holland, 2015). In contrast, in amphioxus, the larval stages with developing gill slits were identified as the least derived. Hence, the least derived stages across the chordates phylum may not be the developmental phase involving patterning of chordate-specific and conserved anatomical structures, as has been implied by previous studies (Holland, 2015; Hu et al., 2017). Alternatively, it is also possible that the phylotypic period of *C. intestinalis* became highly diversified. However, a caveat of our results would be that it may not be suitable to infer relationships between phylogenetic groups and developmental stages [such as which phylogenetic group of animals follow the hourglass model, or to find recapitulative tendencies during development (Kuratani et al., 2021; Uesaka et al., 2021)].

We also pinpointed the turtle showing high transcriptomic derivedness, which is consistent with the morphological perspective, but the result is contrary to those of genomic studies that indicated a slow genomic evolutionary rate (Bradley Shaffer et al., 2013; Green et al., 2014). An even more unexpected finding was that tunicate *C. intestinalis* showed the lowest transcriptomic derivedness indices among all the chordate species compared with recent genomic and morphological studies suggesting that tunicates could be a group of rapid-evolving species (Berná and Alvarez-Valin, 2014; Holland, 2016). Transcriptomic datasets of additional tunicate species may help confirm whether this trend is general to tunicates or specific to *C. intestinalis*. Given that our approach accurately captured transcriptomic derivedness, these results suggest that high transcriptomic derivedness may not directly reflect high evolutionary rates in their genomes inferred from commonly shared, conserved sequences.

Nonetheless, these apparent contradictions highlighted excellent opportunities to further revise and improve our method for evaluating the derivedness of embryos in future studies. Ideally speaking, although derivedness should be evaluated together with species-specific genes, the tree estimated using our method with species-specific genes violated the criterion to show topology consistent with known phylogeny (**Supplementary Figure S4**). This implies that either better criteria for tree topology should be invented or better models of transcriptome evolution should be incorporated when calculating the evolutionary distance between embryonic transcriptomes. Indeed, it could even be possible that transcriptomic derivedness may not perfectly match with phylogenetic relationship deduced from genomic information, breaking the criterion 2 (derivedness tree follows the known phylogeny deduced from genomic information). For example, phylogenetic trees based on genomic sequences mainly rely on those alignable between different species, resulting in exclusion of species-specific genomic sequences. On the other hand, our methodology included information (gene expression levels) that are potentially lost in certain species. These differences could lead to inconsistencies in their phylogenetic tree topologies. In addition, another possibility would be that species having very different evolutionary speed in their phenotypes (including their developmental transcriptome) and genomic mutational rate could also lead to inconsistent phylogenetic results when compared with other species. Second, we proposed assigning “0” expression levels to potentially lost genes. However, in addition to the technical difficulty of definitively identifying lost genes, it also remains unclear how these “0”-expressing ortholog-groups would affect the calculated transcriptomic derivedness, especially species that lost many genes, such as *C. intestinalis*. In addition, further studies are needed to understand how bias can be minimized from other technical aspects, including read depth. For instance, including closely related species, or even populations in the same species, would offer a way to measure technical biases, as these are expected to have similar transcriptomic profiles. Our method could also be potentially biased by the differences in genome annotation quality among species; however, since annotation quality depends on a variety of factors, such as how many genome-sequenced closely related species are available and sequencers or assembly methods used for each species, it is unfeasible, at least at this moment, to adjust the quality of annotations in different species. Importantly, to obtain a more comprehensive picture of the derivedness of embryos, transcriptomes should not be the only parameter to measure; instead, other aspects such as epigenomes, morphologies, and changes observed in fossil records, and cell types could also contain a lot of information about derivedness and should be considered.

Finally, we attempted to characterize the least derived developmental stages of vertebrate species by extracting ortholog-groups showing a negative correlation between the derivedness index and expression levels during development. Remarkably, the extracted orthologous gene set included numerous genes known to be expressed in the shared anatomical structures of vertebrates during this period. Given that the least derived stage may represent a period with the inclination to retain the ancestral phenotype, in line with the recent perspectives that suggested the constraints in this least derived/conserved period may contribute to the strict conservation of animal body plans (Hu et al., 2017; Furusawa and Irie, 2020), this set of orthologous genes (negative DCOs across vertebrate species) may provide additional insights into the evolutionary mechanisms behind the conservative features of body plans. Taken together, further development of the derivedness index could be a useful quantitative indicator to further study which developmental processes are potentially less evolvable [including those argued in line with developmental constraints (Smith et al., 1985; Galis, 1999; Irie and Kuratani, 2014; Hu et al., 2017; Furusawa and Irie, 2020) and developmental burden (Riedl, 1978; Fujimoto et al., 2021)].

4 METHODS

4.1 Animal Use and Care

Experimental procedures and animal care were conducted in strict accordance with the guidelines approved by the University of Tokyo (approval ID: 14-03, 16-2). The animals were sacrificed with minimal suffering. Individual embryos were blindly selected from the wild-type population.

4.2 Embryo Collection, RNA Extraction, Library Preparation, and RNA-Seq

The RNA-seq data utilized were from published datasets (Wang et al., 2013; Hu et al., 2017; Ichikawa et al., 2017; Li et al., 2018, 2020) and three other studies [purple sea urchin (Tu et al., 2012); oyster: (Zhang et al., 2012); *Drosophila*: (Nègre et al., 2011):]. These include major early to-late developmental stages of eight representative chordate species (*Mus musculus*, *Gallus gallus*, *Pelodiscus sinensis*, *Xenopus laevis*, *Danio rerio*, *Oryzias latipes*, *Ciona intestinalis*, *Branchiostoma floridae*), four echinoderms (*Anneissia japonica*, *Apostichopus japonicus*, *Lytechinus variegatus*, and *Strongylocentrotus purpuratus*), and two outgroup species (*Crassostrea gigas*, *Drosophila melanogaster*), with 2-3 biologically independent replicates of each stage to represent the general population (except *S. purpuratus*, *C. gigas*, and *D. melanogaster*). All samples were sequenced using Illumina platforms. Details of the included datasets are summarized in **Supplementary Tables S1–13** (developmental stages covered) and **Supplementary Table S14** (e.g., sample accession numbers, sequencing platforms).

To compare datasets collected by TruSeq and Quartz-Seq (Sasagawa et al., 2013) protocols (the latter involves an additional whole-transcript amplification step), additional samples of mouse E9.0 and E15.5 (C57BL/6J strain, CLEA Japan) were prepared following the same procedures reported in the mouse and sea cucumber datasets (Hu et al., 2017; Li et al., 2018). Both TruSeq and Quartz-seq were performed on the same total RNA extracted from pooled embryos. An average of nine embryos from three independent parents were pooled for each E9.0 sample, and an average of three embryos, E15.0. Library preparation for sequencing (TruSeq: single-end, 100 bp, non-strand-specific; Quartz-Seq: paired-end, 100 bp, non-strand-specific) was performed according to the manufacturer's instructions using the Illumina TruSeq RNA Library Prep Kit v2 or the standard protocol of Quartz-Seq (Sasagawa et al., 2013) with Illumina Nextera XT Library Prep Kit. The normalized libraries were sequenced using an Illumina HiSeq 1,500. The raw reads data are available in SRA through the accession number PRJNA749373.

4.3 RNA-Seq Mapping

Adapter sequence trimming and low-quality read filtering in RNA-seq raw reads were performed with trimmomatic (version 0.36) with default parameters (ILLUMINACLIP:2:30:10, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, and MINLEN:36) (Bolger et al., 2014). Quality checks were performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were aligned to the respective genome using HISAT2 (version 2.1.0) (Kim et al., 2019), with mitochondrial scaffolds and NUMT sequences manually removed from the genomic sequence files to avoid multi-copy quantification biases. The genome versions are summarized in **Supplementary Table S15**. The expression levels of coding genes were quantified in TPM (transcripts per million) (Wagner et al., 2012) using StringTie (version 1.3.4d) (Pertea et al., 2015).

4.4 Construction of Ortholog-Group-Based Expression Tables

Gene-based expression profiles of embryos of all species were summarized into an ortholog-group-based expression table by the method explained in the main text and **Figure 1B**. Our method was modified from the approach previously reported by the EXPANDE Project Consortium (Hu et al., 2017). Genes were first grouped into ortholog-groups using PORTHOMCL (Li et al., 2003; Tabari and Su, 2017) with a BLASTP (version 2.7.1) e-value threshold of 1e-5, as an approach considering paralogs and potentially lost genes. The expression level per ortholog-group was then estimated by the mean (or sum) expression level of all genes in the ortholog-group in each species. The expression level of "0" was assigned to an ortholog-group with no predicted gene in that species.

4.5 Scanning of Methods for the Construction of Derivedness Tree

4.5.1 Derivedness Tree Construction

The derivation of the derivedness tree was based on a pairwise evaluation of transcriptomic similarities between embryos. The tree was inferred from the calculated distance matrix based on the ortholog-group-based expression table (**Figure 2A**). The most suitable method was searched through combinations of major methods for expression data normalization, distance calculation, and tree inference (**Table 1**) to meet the three criteria discussed in the main text (**Figure 2B**). In expression data normalization, log (TPM+1) was used to avoid the undefined value of "log0." The trees were plotted using the "ggtree" package in R (Yu et al., 2016).

4.5.2 Biological Replicates of Expression Data

For each developmental stage, the expression level of a gene can be represented by the mean value of its expression levels in all biological replicates. These mean-value expression data were used to generate the tree shown in **Figure 3A**, and all other trees in the **Supplementary Material**, if not further specified. However, this method may incorporate false-positives and false-negatives in individual sample sets. To avoid this potential bias, we incorporated deviations of gene expression levels in different biological replicates and created "biological replicates-included (BRI)" expression data: a set of BRI expression data randomly takes one biological-replicate sample for each developmental stage. As such, many combinations can be acquired (for example, three biological replicates for 10 developmental stages each can create 310 different combinations) to simulate expression changes during early-to-late development. Support values on the tree in **Figure 3A** were calculated as the consensus of trees inferred from 100 random BRI expression tables, and these values represent the strength of the biological replicates supporting the tree topology. For the trees inferred from Spearman distance, the consensus tree of 100 BRI trees showed a topology similar to that inferred from the mean-value expression data.

4.5.3 Smoothness Analysis

As gene expression profiles are expected to show gradual changes during development, this is reflected in the distance matrix showing a smooth texture. We adapted texture analysis in image processing theory (Materka and Strzelecki, 1998; Gonzalez and Woods, 2007) to investigate which distance method generated the smoothest matrix image in the within-species comparison regions (**Figures 2B–3–3**). The smoothness of each within-species comparison region on the distance matrix image was measured using the following descriptor statistics: homogeneity, dissimilarity, contrast, uniformity, and correlation (**Supplementary Table S16**), with a sliding window of 3 × 3 using the GLCM package in Python. For each within-species region on the distance matrix image, the distance method that scored the best according to each descriptor was recorded, and the percentage of each distance method being selected as the best-scoring method is plotted in **Figure 2D**.

4.6 Analyses Using Derivedness Index

The branch length of each embryo from the common ancestral node on the tree (putative common ancestor of chordates and echinoderms) was extracted and defined as the derivedness index. In calculating derivedness indices for species as a whole (for whole life cycles), the range of derivedness indices of all the developmental stages in the inferred tree (**Figure 3A**) was represented as box plots (whether the differences among species were statistically significant were examined by Kruskal-Wallis test; **Figure 3B**). To compare the derivedness indices of penta-radial phase against pre-metamorphosis stages in each echinoderm species (**Figure 3D**), the range of the derivedness indices of the corresponding developmental phase was represented as a box plot (Mann-Whitney *U* test). Penta-radial phase for each species was defined as: (feather star) doliolaria onwards; (sea cucumber) metamorph-1 onwards; (green sea urchin) 7wpf onwards; (purple sea urchin) vestibular-invagi onwards. To examine the derivedness index of each developmental stage (**Figure 4**), the results from the 100 random BRI trees were utilized, and for each developmental stage, the range of derivedness indices in the 100 trees was summarized for each embryo (whether the differences among developmental stages were statistically significant were examined by Friedman test).

4.7 Identification of Derivedness-Correlative Ortholog-Groups

Samples collected by Quartz-Seq were omitted from the analysis to avoid any unexpected bias. For each species and ortholog-group, Spearman's correlation coefficient between its expression levels and derivedness indices along early to late embryos was calculated. Ortholog-groups showing negative correlations across certain groups of animals (vertebrates, chordates, and echinoderms) were further analyzed. A total of 7,775 DCOs, with at least one gene counterpart in each vertebrate species, were classified into seven categories based on the number of vertebrate species that showed a negative correlation.

4.7.1 Gene Name and Functional Prediction of Ortholog-Groups

To identify the names of genes included in each ortholog-group, predictions from two sources were incorporated, namely, the genome annotation file and names predicted by PANNZER2 (Törönen et al., 2018). First, the name of each gene, if any, was extracted from the genome annotation files. For prediction by PANNZER2, the peptide sequence of the longest isoform of each gene was used, and the predicted gene names and GO terms were retrieved. Predicted GO terms were mapped to GOslim terms (go.obo release version 2020–03–23) using GOATOOLS (Klopfenstein et al., 2018). Development-related ortholog-groups in **Figure 5C** were defined as those including genes with the GOslim term of "anatomical structure development" (GO:0048856). For Hox ortholog-groups (**Figure 5C**, left), predictions from both sources were manually checked because not every single *Hox* gene was conserved in all species.

4.7.2 Tendency of Temporal Pleiotropy

For each category of vertebrate DCOs, the proportion of ortholog-groups showing different degrees of temporal pleiotropy across species was calculated. Since the number of sampled developmental stages was not uniform in different species, for each ortholog-group, the percentage of developmental stages where it is expressed (defined as TPM ≥1) in each species was first calculated. The mean percentage across the six vertebrate species was then calculated for each ortholog-group, and the whole range of mean percentage values was binned into ten 10% ranges to be plotted. For each category, the ratio of ortholog-groups with a mean percentage value within each bin is shown in **Figure 5E**. In other words, ortholog-groups inside the 100% bin (blue) indicate that they are expressed in all the sampled developmental stages in all six vertebrate species. Similarly, those inside the 90% bin (light blue) are expressed in 90–99% of all the sampled developmental stages on average across species. Thus, ortholog-groups showing a higher degree of temporal pleiotropy are shown towards the right end, while those with a lower degree of pleiotropy are towards the left side.

4.8 Read-Depth Normalization Between Species

To normalize read depth between samples of different species, the maximum possible number of reads was selected while keeping the same reads-to-exome-size ratio in all species. Best-hits were first selected using samtools (version 1.12; single-end samples: samtools view -F 4 | samtools view -F 256; paired-end samples: samtools view -f 2 | samtools view -F 256) (Li et al., 2009). Random pick-up of reads was performed from the BED file containing all mapped best-hit reads. For paired-end samples, only one read was reported for the BED file using the bamtobed function with the "-bedpe" option from bedtools (version 2.30.0) (Quinlan and Hall, 2010). The ratio of all best-hit reads to exome size in each sample was calculated, and the smallest ratio was selected as the baseline (Lv_60cell sample) to calculate the number of reads needed for each species (see **Supplementary Table S18** for the calculated results). We tried to calculate exome size using the original genome annotation file as well as a version with all the untranslated regions (UTRs) removed to minimize unintentional bias, since only some of the species were annotated with UTRs (**Supplementary Figure S15**). To remove UTRs from annotation files, the script "gff3_file_UTR_trimmer.pl" from PASApipeline (Haas et al., 2003) was used. To calculate exome size, the command line "awk '{if ($3 = = "exon"){print $0}}' $gffname | gff2bed - | bedops -m - | awk 'BEGIN{FS = "\t"; count = 0}{count = count + ($3-$2)}END{print count}'" with the BEDOPS tool (version 2.4.39) (Neph et al., 2012) was used.

4.9 Software and Computation Environment

Bioinformatics analyses were performed using in-house R (4.0.3) (R-Core-Team, 2020), Python (3.7), and shell scripts, together with the software and packages summarized in **Supplementary Table S17**.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA749373.

ETHICS STATEMENT

The animal study was reviewed and approved by The University of Tokyo.

AUTHOR CONTRIBUTIONS

The study was conceived and supervised by NI Bioinformatic analyses were performed by JL and NI Genetic resources were contributed by all authors.

FUNDING

This work was supported by the National Natural Science Foundation of China (No.31561143016; No. 31621062) and the Strategic International Collaborative Research Program (SICORP) of JST.

ACKNOWLEDGMENTS

We thank Flore Castellan for providing illustrations of the echinoderm embryos. We thank one of the reviewers for providing a specific idea to further test potential technical biases that was discussed in the Discussion section.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.749963/full#supplementary-material

Supplementary Table S19 | 695 DCOs (derivedness-correlative ortholog-groups) showing negative correlations across six vertebrate species.

Supplementary Table S20 | 230 negative DCOs across eight chordate species.

Supplementary Table S21 | 2,414 negative DCOs across three echinoderm species.

REFERENCES

Amemiya, C. T., Alföldi, J., Lee, A. P., Fan, S., Philippe, H., MacCallum, I., et al. (2013). The African Coelacanth Genome Provides Insights into Tetrapod Evolution. *Nature* 496, 311–316. doi:10.1038/nature12027

Andrioli, L. P. (2012). Toward New Drosophila Paradigms. *Genesis* 50, 585–598. doi:10.1002/dvg.22019

Arthur, W. (1997). *The Origin of Animal Body Plans* Cambridge, UK: Cambridge University Press. doi:10.1017/cbo9781139174596

Ashique, A. M., Choe, Y., Karlen, M., May, S. R., Phamluong, K., Solloway, M. J., et al. (2009). The Rfx4 Transcription Factor Modulates Shh Signaling by Regional Control of Ciliogenesis. *Sci. Signal.* 2, ra70. doi:10.1126/scisignal.2000602

Baer, C. F., Miyamoto, M. M., and Denver, D. R. (2007). Mutation Rate Variation in Multicellular Eukaryotes: Causes and Consequences. *Nat. Rev. Genet.* 8, 619–631. doi:10.1038/nrg2158

Baum, D. A., and Smith, S. D. (2012). *Tree Thinking*. Greenwood Village, CO, USA: Roberts and Company Publishers, Inc..

Belgacem, M. R., Escande, M.-l., Escriva, H., and Bertrand, S. (2011). Amphioxus Tbx6/16 and Tbx20 Embryonic Expression Patterns Reveal Ancestral Functions in Chordates. *Gene Expr. Patterns* 11, 239–243. doi:10.1016/j.gep.2010.12.006

Benton, M. J. (2014). *Vertebrate Palaeontology*. 4th ed. Chichester, UK: Wiley-Blackwell.

Berná, L., and Alvarez-Valin, F. (2014). Evolutionary Genomics of Fast Evolving Tunicates. *Genome Biol. Evol.* 6, 1724–1738. doi:10.1093/gbe/evu122

Bogdanović, O., Smits, A. H., de la Calle Mustienes, E., Tena, J. J., Ford, E., Williams, R., et al. (2016). Active DNA Demethylation at Enhancers during the Vertebrate Phylotypic Period. *Nat. Genet.* 48, 417–426. doi:10.1038/ng.3522

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170

Bowes, J. B., Snyder, K. A., Segerdell, E., Jarabek, C. J., Azam, K., Zorn, A. M., et al. (2010). Xenbase: Gene Expression and Improved Integration. *Nucleic Acids Res.* 38, D607–D612. doi:10.1093/nar/gkp953

Bradley Shaffer, H., Minx, P., Warren, D. E., Shedlock, A. M., Thomson, R. C., Valenzuela, N., et al. (2013). The Western Painted Turtle Genome, a Model for the Evolution of Extreme Physiological Adaptations in a Slowly Evolving Lineage. *Genome Biol.* 14, R28. doi:10.1186/gb-2013-14-3-r28

Bromham, L., and Penny, D. (2003). The Modern Molecular Clock. *Nat. Rev. Genet.* 4, 216–224. doi:10.1038/nrg1020

Brusca, R. C., Moore, W., and Shuster, S. M. (2016). *Invertebrates*. 3rd ed. Sunderland, MA, U.S.A.: Sinauer Associates.

Cannon, J. T., Kocot, K. M., Waits, D. S., Weese, D. A., Swalla, B. J., Santos, S. R., et al. (2014). Phylogenomic Resolution of the Hemichordate and Echinoderm Clade. *Curr. Biol.* 24, 2827–2832. doi:10.1016/j.cub.2014.10.016

Catela, C., Shin, M. M., and Dasen, J. S. (2015). Assembly and Function of Spinal Circuits for Motor Control. *Annu. Rev. Cel Dev. Biol.* 31, 669–698. doi:10.1146/annurev-cellbio-100814-125155

Darnell, D. K., Kaur, S., Stanislaw, S., Davey, S., Konieczka, J. H., Yatskievych, T. A., et al. (2007). GEISHA: an *In Situ* Hybridization Gene Expression Resource for the Chicken Embryo. *Cytogenet. Genome Res.* 117, 30–35. doi:10.1159/000103162

Diogo, R., Ziermann, J. M., and Linde-Medina, M. (2015). Is Evolutionary Biology Becoming Too Politically Correct? A Reflection Onthe Scala Naturae, Phylogenetically Basal Clades, Anatomically Plesiomorphic Taxa, and 'lower' Animals. *Biol. Rev.* 90, 502–521. doi:10.1111/brv.12121

Domazet-Lošo, T., and Tautz, D. (2010). A Phylogenetically Based Transcriptome Age index Mirrors Ontogenetic Divergence Patterns. *Nature* 468, 815–818. doi:10.1038/nature09632

Duboule, D. (1994). Temporal Colinearity and the Phylotypic Progression: A Basis for the Stability of a Vertebrate Bauplan and the Evolution of Morphologies through Heterochrony. *Development*, 135–142. doi:10.1242/dev.1994.supplement.135

Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., and Peterson, K. J. (2011). The Cambrian Conundrum: Early Divergence and Later Ecological success in the Early History of Animals. *Science* 334, 1091–1097. doi:10.1126/science.1206375

Fitch, W. M., and Margoliash, E. (1967). Construction of Phylogenetic Trees. *Science* 155, 279–284. doi:10.1126/science.155.3760.279

Fujimoto, S., Yamanaka, K., Tanegashima, C., Nishimura, O., Kuraku, S., Kuratani, S., et al. (2021). Measuring Potential Effects of the Developmental burden Associated with the Vertebrate Notochord. *J. Exp. Zool B Mol. Dev. Evol.* doi:10.1002/jez.b.23032

Furusawa, C., and Irie, N. (2020). Toward Understanding of Evolutionary Constraints: Experimental and Theoretical Approaches. *Biophys. Rev.* 12, 1155–1161. doi:10.1007/s12551-020-00708-2

Galis, F. (1999). Why Do Almost All Mammals Have Seven Cervical Vertebrae? Developmental constraints,Hox Genes, and Cancer. *J. Exp. Zool.* 285, 19–26. doi:10.1002/(sici)1097-010x(19990415)285:1<19:aid-jez3>3.0.co;2-z

Gascuel, O. (1997). BIONJ: an Improved Version of the NJ Algorithm Based on a Simple Model of Sequence Data. *Mol. Biol. Evol.* 14, 685–695. doi:10.1093/oxfordjournals.molbev.a025808

Gaut, B., Yang, L., Takuno, S., and Eguiarte, L. E. (2011). The Patterns and Causes of Variation in Plant Nucleotide Substitution Rates. *Annu. Rev. Ecol. Evol. Syst.* 42, 245–266. doi:10.1146/annurev-ecolsys-102710-145119

Gemmell, N. J., Rutherford, K., Prost, S., Tollis, M., Winter, D., Macey, J. R., et al. (2020). The Tuatara Genome Reveals Ancient Features of Amniote Evolution. *Nature* 584, 403–409. doi:10.1038/s41586-020-2561-9

Gilbert, S. F., and Corfe, I. (2013). Turtle Origins: Picking up Speed. *Dev. Cel* 25, 326–328. doi:10.1016/j.devcel.2013.05.011

Gonzalez, R. C., and Woods, R. E. (2007). "Regional Descriptors: Texture," in *Digital Image Processing*. Upper Saddle Rier, NJ: Pearson Prentice Hall, 827–839.

Goodrich, L. V., Jung, D., Higgins, K. M., and Scott, M. P. (1999). Overexpression of Ptc1 Inhibits Induction of Shh Target Genes and Prevents normal Patterning in the Neural Tube. *Dev. Biol.* 211, 323–334. doi:10.1006/dbio.1999.9311

Green, R. E., Braun, E. L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., et al. (2014). Three Crocodilian Genomes Reveal Ancestral Patterns of Evolution Among Archosaurs. *Science* 346, 1254449. doi:10.1126/science.1254449

Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Jr, R. K. S., Hannick, L. I., et al. (2003). Improving the Arabidopsis Genome Annotation Using Maximal Transcript Alignment Assemblies. *Nucleic Acids Res.* 31, 5654–5666. doi:10.1093/nar/gkg770

Hazkani-Covo, E., Wool, D., and Graur, D. (2005). In search of the vertebrate phylotypic stage: A molecular examination of the developmental hourglass model and von Baer's third law. *J. Exp. Zool.* 304B, 150–158. doi:10.1002/jez.b.21033

Hoekstra, H. E. (2006). Genetics, Development and Evolution of Adaptive Pigmentation in Vertebrates. *Heredity* 97, 222–234. doi:10.1038/sj.hdy.6800861

Hogan, J. D., Keenan, J. L., Luo, L., Ibn-Salem, J., Lamba, A., Schatzberg, D., et al. (2020). The Developmental Transcriptome for Lytechinus variegatus Exhibits Temporally Punctuated Gene Expression Changes. *Dev. Biol.* 460, 139–154. doi:10.1016/j.ydbio.2019.12.002

Holland, L. Z. (2015). Genomics, Evolution and Development of Amphioxus and Tunicates: The Goldilocks Principle. *J. Exp. Zool. (Mol. Dev. Evol.* 324, 342–352. doi:10.1002/jez.b.22569

Holland, L. Z. (2016). Tunicates. *Curr. Biol.* 26, R146–R152. doi:10.1016/j.cub.2015.12.024

Hu, H., Uesaka, M., Uesaka, M., Guo, S., Shimai, K., Lu, T.-M., et al. (2017). Constrained Vertebrate Evolution by Pleiotropic Genes. *Nat. Ecol. Evol.* 1, 1722–1730. doi:10.1038/s41559-017-0318-0

Hyman, L. H. (1955). *Echinodermata: The Coelomate Bilateria*. New York: McGraw-Hill.

Ichikawa, K., Tomioka, S., Suzuki, Y., Nakamura, R., Doi, K., Yoshimura, J., et al. (2017). Centromere Evolution and CpG Methylation during Vertebrate Speciation. *Nat. Commun.* 8, 1833. doi:10.1038/s41467-017-01982-7

Irie, N., and Kuratani, S. (2011). Comparative Transcriptome Analysis Reveals Vertebrate Phylotypic Period during Organogenesis. *Nat. Commun.* 2, 248. doi:10.1038/ncomms1248

Irie, N., and Kuratani, S. (2014). The Developmental Hourglass Model: a Predictor of the Basic Body Plan? *Development* 141, 4649–4655. doi:10.1242/dev.107318

Irie, N., Satoh, N., and Kuratani, S. (2018). The Phylum Vertebrata: a Case for Zoological Recognition. *Zoolog. Lett* 4, 32–20. doi:10.1186/s40851-018-0114-y

Irie, N., and Sehara-Fujisawa, A. (2007). The Vertebrate Phylotypic Stage and an Early Bilaterian-Related Stage in Mouse Embryogenesis Defined by Genomic Information. *BMC Biol.* 5, 1. doi:10.1186/1741-7007-5-1

Kalinka, A. T., Varga, K. M., Gerrard, D. T., Preibisch, S., Corcoran, D. L., Jarrells, J., et al. (2010). Gene Expression Divergence Recapitulates the Developmental Hourglass Model. *Nature* 468, 811–814. doi:10.1038/nature09634

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* 37, 907–915. doi:10.1038/s41587-019-0201-4

Klopfenstein, D. V., Zhang, L., Pedersen, B. S., Ramírez, F., Warwick Vesztrocy, A., Naldi, A., et al. (2018). GOATOOLS: a Python Library for Gene Ontology Analyses. *Sci. Rep.* 8, 10872–10917. doi:10.1038/s41598-018-28948-z

Kuratani, S., Uesaka, M., and Irie, N. (2021). How Can Recapitulation Be Reconciled with Modern Concepts of Evolution?. *J. Exp. Zool (Mol Dev. Evol.* doi:10.1002/jez.b.23020

Lanfear, R., Welch, J. J., and Bromham, L. (2010). Watching the Clock: Studying Variation in Rates of Molecular Evolution between Species. *Trends Ecol. Evol.* 25, 495–503. doi:10.1016/j.tree.2010.06.007

Lefort, V., Desper, R., and Gascuel, O. (2015). FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program: Table 1. *Mol. Biol. Evol.* 32, 2798–2800. doi:10.1093/molbev/msv150

Leo, L., Yu, W., D'Rozario, M., Waddell, E. A., Marenda, D. R., Baird, M. A., et al. (2015). Vertebrate Fidgetin Restrains Axonal Growth by Severing Labile Domains of Microtubules. *Cel Rep.* 12, 1723–1730. doi:10.1016/j.celrep.2015.08.017

Levin, M., Hashimshony, T., Wagner, F., and Yanai, I. (2012). Developmental Milestones Punctuate Gene Expression in the Caenorhabditis Embryo. *Dev. Cel* 22, 1101–1108. doi:10.1016/j.devcel.2012.04.004

Li, C., Wu, X.-C., Rieppel, O., Wang, L.-T., and Zhao, L.-J. (2008). An Ancestral Turtle from the Late Triassic of Southwestern China. *Nature* 456, 497–501. doi:10.1038/nature07533

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

Li, L., Stoeckert, C. J., Jr, and Roos, D. S. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res.* 13, 2178–2189. doi:10.1101/gr.1224503

Li, W.-H., Ellsworth, D. L., Krushkal, J., Chang, B. H.-J., and Hewett-Emmett, D. (1996). Rates of Nucleotide Substitution in Primates and Rodents and the Generation-Time Effect Hypothesis. *Mol. Phylogenet. Evol.* 5, 182–187. doi:10.1006/mpev.1996.0012

Li, Y., Kikuchi, M., Li, X., Gao, Q., Xiong, Z., Ren, Y., et al. (2018). Weighted Gene Co-expression Network Analysis Reveals Potential Genes Involved in Early Metamorphosis Process in Sea Cucumber Apostichopus Japonicus. *Biochem. Biophysical Res. Commun.* 495, 1395–1402. doi:10.1016/j.bbrc.2017.11.154

Li, Y., Omori, A., Flores, R. L., Satterfield, S., Nguyen, C., Ota, T., et al. (2020). Genomic Insights of Body Plan Transitions from Bilateral to Pentameral Symmetry in Echinoderms. *Commun. Biol.* 3, 371. doi:10.1038/s42003-020-1091-1

Lynch, M., Ackerman, M. S., Gout, J.-F., Long, H., Sung, W., Thomas, W. K., et al. (2016). Genetic Drift, Selection and the Evolution of the Mutation Rate. *Nat. Rev. Genet.* 17, 704–714. doi:10.1038/nrg.2016.104

Markow, T. A., and O'Grady, P. M. (2007). Drosophila Biology in the Genomic Age. *Genetics* 177, 1269–1276. doi:10.1534/genetics.107.074112

Marlétaz, F., Firbas, P. N., Maeso, I., Tena, J. J., Bogdanovic, O., Perry, M., et al. (2018). Amphioxus Functional Genomics and the Origins of Vertebrate Gene Regulation. *Nature* 564, 64–70. doi:10.1038/s41586-018-0734-6

Martin, A. P., and Palumbi, S. R. (1993). Body Size, Metabolic Rate, Generation Time, and the Molecular Clock. *Proc. Natl. Acad. Sci.* 90, 4087–4091. doi:10.1073/pnas.90.9.4087

Materka, A., and Strzelecki, M. (1998). Texture Analysis Methods – a Review. COST B11 Technical Report. Lodz-Brussels: Technical University of Lodz. Available at: http://www.eletel.p.lodz.pl/programy/cost/pdf_1.pdf (Accessed June 25, 2021).

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA, USA: Belknap Press.

Miller, H. C., Moore, J. A., Allendorf, F. W., and Daugherty, C. H. (2009). The Evolutionary Rate of Tuatara Revisited. *Trends Genet.* 25, 13–15. doi:10.1016/j.tig.2008.09.007

Morris, V. B. (1999). Bilateral Homologues in Echinoderms and a Predictive Model of the Bilateral Echinoderm Ancestor. *Biol. J. Linn. Soc.* 66, 293–303. doi:10.1111/j.1095-8312.1999.tb01892.x

Murdoch, J. N., and Copp, A. J. (2010). The Relationship between Sonic Hedgehog Signaling, Cilia, and Neural Tube Defects. *Birth Defects Res. A: Clin. Mol. Teratology* 88, 633–652. doi:10.1002/bdra.20686

Nagashima, H., Sugahara, F., Takechi, M., Ericsson, R., Kawashima-Ohya, Y., Narita, Y., et al. (2009). Evolution of the Turtle Body Plan by the Folding and Creation of New Muscle Connections. *Science* 325, 193–196. doi:10.1126/science.1173826

Nègre, N., Brown, C. D., Ma, L., Bristow, C. A., Miller, S. W., Wagner, U., et al. (2011). A Cis-Regulatory Map of the Drosophila Genome. *Nature* 471, 527–531. doi:10.1038/nature09990

Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., et al. (2012). BEDOPS: High-Performance Genomic Feature Operations. *Bioinformatics* 28, 1919–1920. doi:10.1093/bioinformatics/bts277

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie Enables Improved Reconstruction of a Transcriptome from RNA-Seq Reads. *Nat. Biotechnol.* 33, 290–295. doi:10.1038/nbt.3122

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* 26, 841–842. doi:10.1093/bioinformatics/btq033

R-Core-Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Revell, L. J., and Chamberlain, S. A. (2014). Rphylip: an R Interface for PHYLIP. *Methods Ecol. Evol.* 5, 976–981. doi:10.1111/2041-210x.12233

Richardson, L., Venkataraman, S., Stevenson, P., Yang, Y., Moss, J., Graham, L., et al. (2014). EMAGE Mouse Embryo Spatial Gene Expression Database: 2014 Update. *Nucl. Acids Res.* 42, D835–D844. doi:10.1093/nar/gkt1155

Riedl, R. (1978). *Order in Living Organisms: Systems Analysis of Evolution*. Chichester, NY, USA: John Wiley & Sons.

Saitou, N., and Nei, M. (1987). The Neighbor-Joining Method: a New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* 4 (4), 406–425. doi:10.1093/oxfordjournals.molbev.a040454

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K. D., Imai, T., et al. (2013). Quartz-Seq: a Highly Reproducible and Sensitive Single-Cell RNA Sequencing Method, Reveals Non-genetic Gene-Expression Heterogeneity. *Genome Biol.* 14, R31. doi:10.1186/gb-2013-14-4-r31

Sedykh, I., Keller, A. N., Yoon, B., Roberson, L., Moskvin, O. V., and Grinblat, Y. (2018). Zebrafish Rfx4 Controls Dorsal and Ventral Midline Formation in the Neural Tube. *Dev. Dyn.* 247, 650–659. doi:10.1002/dvdy.24613

Smith, C. M., Hayamizu, T. F., Finger, J. H., Bello, S. M., McCright, I. J., Xu, J., et al. (2018). The Mouse Gene Expression Database (GXD): 2019 Update. *Nucleic Acids Res.* 47, D774–D779. doi:10.1093/nar/gky922

Smith, J. M., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., et al. (1985). Developmental Constraints and Evolution: a Perspective from the mountain lake Conference on Development and Evolution. *Q. Rev. Biol.* 60, 265–287. doi:10.1086/414425

Stern, D. (2001). Body-size Evolution: How to Evolve a mammoth Moth. *Curr. Biol.* 11, R917–R919. doi:10.1016/s0960-9822(01)00554-1

Tabari, E., and Su, Z. (2017). PorthoMCL: Parallel Orthology Prediction Using MCL for the Realm of Massive Genome Availability. *Big Data Anal.* 2, 4. doi:10.1186/s41044-016-0019-8

Thisse, B., Heyer, V., Lux, A., Alunni, V., Degrave, A., Seiliez, I., et al. (2004). Spatial and Temporal Expression of the Zebrafish Genome by Large-Scale *In Situ* Hybridization Screening. *Methods Cel Biol* 77, 505–519. doi:10.1016/s0091-679x(04)77027-2

Törönen, P., Medlar, A., and Holm, L. (2018). Pannzer2: a Rapid Functional Annotation Web Server. *Nucleic Acids Res.* 46, W84–W88. doi:10.1093/nar/gky350

Tu, Q., Cameron, R. A., Worley, K. C., Gibbs, R. A., and Davidson, E. H. (2012). Gene Structure in the Sea Urchin Strongylocentrotus purpuratus Based on Transcriptome Analysis. *Genome Res.* 22, 2079–2087. doi:10.1101/gr.139170.112

Uchida, Y., Uesaka, M., Yamamoto, T., Takeda, H., and Irie, N. (2018). Embryonic Lethality Is Not Sufficient to Explain Hourglass-like Conservation of Vertebrate Embryos. *Evodevo* 9, 7. doi:10.1186/s13227-018-0095-0

Uesaka, M., Kuratani, S., and Irie, N. (2021). The Developmental Hourglass Model and Recapitulation: An Attempt to Integrate the Two Models. *J. Exp. Zool (Mol Dev. Evol.* doi:10.1002/jez.b.23027

Uesaka, M., Kuratani, S., Takeda, H., and Irie, N. (2019). Recapitulation-like Developmental Transitions of Chromatin Accessibility in Vertebrates. *Zoolog. Lett* 5, 33–15. doi:10.1186/s40851-019-0148-9

Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., and Reece, J. B. (2016). *Campbell Biology*. New York, NY, USA: Pearson, 560.

Wagner, G. P., Kin, K., and Lynch, V. J. (2012). Measurement of mRNA Abundance Using RNA-Seq Data: RPKM Measure Is Inconsistent Among Samples. *Theor. Biosci.* 131, 281–285. doi:10.1007/s12064-012-0162-3

Wang, Z., Pascual-Anaya, J., Zadissa, A., Li, W., Niimura, Y., Huang, Z., Li, C., White, S., Xiong, Z., Fang, D., Wang, B., Ming, Y., Chen, Y., Zheng, Y., Kuraku, S., Pignatelli, M., Herrero, J., Beal, K., Nozawa, M., Li, Q., Wang, J., Zhang, H., Yu, L., Shigenobu, S., Wang, J., Liu, J., Flicek, P., Searle, S., Wang, J., Kuratani, S., Yin, Y., Aken, B., Zhang, G., and Irie, N. (2013). The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and Evolution of the Turtle-specific Body Plan. *Nat. Genet.* 45, 701–706. doi:10.1038/ng.2615

Woolfit, M. (2009). Effective Population Size and the Rate and Pattern of Nucleotide Substitutions. *Biol. Lett.* 5, 417–420. doi:10.1098/rsbl.2009.0155

Xu, F., Domazet-Lošo, T., Fan, D., Dunwell, T. L., Li, L., Fang, X., et al. (2016). High Expression of New Genes in Trochophore Enlightening the Ontogeny and Evolution of Trochozoans. *Sci. Rep.* 6, 34664. doi:10.1038/srep34664

Yokoyama, S., Ito, Y., Ueno-Kudoh, H., Shimizu, H., Uchibe, K., Albini, S., et al. (2009). A Systems Approach Reveals that the Myogenesis Genome Network Is Regulated by the Transcriptional Repressor RP58. *Dev. Cel* 17, 836–848. doi:10.1016/j.devcel.2009.10.011

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T. Y. (2016). Ggtree : an R Package for Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data. *Methods Ecol. Evol.* 8, 28–36. doi:10.1111/2041-210x.12628

Zalts, H., and Yanai, I. (2017). Developmental Constraints Shape the Evolution of the Nematode Mid-developmental Transition. *Nat. Ecol. Evol.* 1, 113. doi:10.1038/s41559-017-0113

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., et al. (2012). The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation. *Nature* 490, 49–54. doi:10.1038/nature11413

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Leong, Li, Uesaka, Uchida, Omori, Hao, Wan, Dong, Ren, Zhang, Zeng, Wang, Chen, Wessel, Livingston, Bradham, Wang and Irie. 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) and the copyright owner(s) 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 Cell and Developmental Biology

METHODS
published: 10 February 2022
doi: 10.3389/fcell.2022.750833

Efficient CRISPR Mutagenesis in Sturgeon Demonstrates Its Utility in Large, Slow-Maturing Vertebrates

Jan Stundl[1,2,3], **Vladimír Soukup**[1], **Roman Franěk**[3], **Anna Pospisilova**[1], **Viktorie Psutkova**[1], **Martin Pšenička**[3], **Robert Cerny**[1], **Marianne E. Bronner**[2], **Daniel Meulemans Medeiros**[4] and **David Jandzik**[1,5]*

[1]Department of Zoology, Faculty of Science, Charles University, Prague, Czechia, [2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, United States, [3]South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Faculty of Fisheries and Protection of Waters, University of South Bohemia in České Budějovice, Vodňany, Czechia, [4]Department of Ecology and Evolutionary Biology, University of Colorado in Boulder, Boulder, CO, United States, [5]Department of Zoology, Faculty of Natural Sciences, Comenius University in Bratislava, Bratislava, Slovakia

In the last decade, the CRISPR/Cas9 bacterial virus defense system has been adapted as a user-friendly, efficient, and precise method for targeted mutagenesis in eukaryotes. Though CRISPR/Cas9 has proven effective in a diverse range of organisms, it is still most often used to create mutant lines in lab-reared genetic model systems. However, one major advantage of CRISPR/Cas9 mutagenesis over previous gene targeting approaches is that its high efficiency allows the immediate generation of near-null mosaic mutants. This feature could potentially allow genotype to be linked to phenotype in organisms with life histories that preclude the establishment of purebred genetic lines; a group that includes the vast majority of vertebrate species. Of particular interest to scholars of early vertebrate evolution are several long-lived and slow-maturing fishes that diverged from two dominant modern lineages, teleosts and tetrapods, in the Ordovician, or before. These early-diverging or "basal" vertebrates include the jawless cyclostomes, cartilaginous fishes, and various non-teleost ray-finned fishes. In addition to occupying critical phylogenetic positions, these groups possess combinations of derived and ancestral features not seen in conventional model vertebrates, and thus provide an opportunity for understanding the genetic bases of such traits. Here we report successful use of CRISPR/Cas9 mutagenesis in one such non-teleost fish, sterlet *Acipenser ruthenus*, a small species of sturgeon. We introduced mutations into the genes *Tyrosinase*, which is needed for melanin production, and *Sonic hedgehog*, a pleiotropic developmental regulator with diverse roles in early embryonic patterning and organogenesis. We observed disruption of both loci and the production of consistent phenotypes, including both near-null mutants' various hypomorphs. Based on these results, and previous work in lamprey and amphibians, we discuss how CRISPR/Cas9 F0 mutagenesis may be successfully adapted to other long-lived, slow-maturing aquatic vertebrates and identify the ease of obtaining and injecting eggs and/or zygotes as the main challenges.

Keywords: CRISPR/Cas9, targeted mutagenesis, non-teleost fish, sturgeon, vertebrates, development, evolution, evo-devo

OPEN ACCESS

Edited by:
Juan Pascual-Anaya,
Malaga University, Spain

Reviewed by:
Tetsuya Nakamura Rutgers,
The State University of New Jersey,
United States
Johanna Kowalko,
Lehigh University, United States

***Correspondence:**
David Jandzik
David.Jandzik@uniba.sk

Specialty section:
This article was submitted to
Evolutionary Developmental Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 31 July 2021
Accepted: 17 January 2022
Published: 10 February 2022

Citation:
Stundl J, Soukup V, Franěk R,
Pospisilova A, Psutkova V,
Pšenička M, Cerny R, Bronner ME,
Medeiros DM and Jandzik D (2022)
Efficient CRISPR Mutagenesis in
Sturgeon Demonstrates Its Utility in
Large, Slow-Maturing Vertebrates.
Front. Cell Dev. Biol. 10:750833.
doi: 10.3389/fcell.2022.750833

INTRODUCTION

The central problem in modern biology is understanding how an organism's one-dimensional genotype, i.e., its linear sequence of nucleotides, gives rise to its four-dimensional phenotype, i.e., its form and function through space and time. For decades, genotype was linked to phenotype by mapping genetic lesions in purebred lines of genetic model organisms. These organisms were carefully selected based on their rapid life cycles and ability to be easily maintained in the laboratory (Hedges, 2002). Mutations occurred either naturally, or were introduced randomly in the genome using radiation, chemicals, or transposable elements. In the 1980s and 1990s, various methods were developed that allowed targeted mutagenesis in select model organisms including *Drosophila*, *Caenorhabditis elegans*, and mouse (Robertson et al., 1988; Kaiser and Goodwin, 1990; Jansen et al., 1997; Faruqi et al., 1998). In the 2000s new "one-size-fits-all" gene targeting technologies, including TALENs and Zinc finger nucleases, allowed targeted mutagenesis in a greater variety of genetic models, such as zebrafish (Doyon et al., 2008; Meng et al., 2008; Huang et al., 2011; Wood et al., 2011). While effective, these methods still required the generation of purebred lines to determine the complete phenotype caused by the mutation. Thus, while substantially faster and more refined than previous methods, gene targeting was still largely limited to a handful of conventional genetic model systems. Furthermore, because these organisms were chosen specifically for their atypical life histories, our understanding of gene function was still largely limited to a few isolated twigs on the vast tree of life.

In the 2010s, these gene targeting strategies became largely supplanted by CRISPR/Cas9-mediated mutagenesis (Bassett et al., 2013; Chen et al., 2013; Fu et al., 2013; Hwang et al., 2013; Li-En et al., 2013; Nakayama et al., 2013; Waaijers et al., 2013; Port et al., 2014; Square et al., 2015; Véron et al., 2015; Martin et al., 2016; Suzuki et al., 2018; Rasys et al., 2019). The CRISPR/Cas9 method results in small, targeted lesions when the cell's DNA repair mechanisms respond to double-stranded DNA breaks created by the Cas9 endonuclease (Barrangou et al., 2007; Jinek et al., 2012; Mali et al., 2013). Practically, CRISPR/Cas9 mutagenesis has several distinct advantages over TALENs and Zinc finger nucleases. Most significantly is its ease of use (Ran et al., 2013). Cas9 endonucleases are commercially available, and quickly and affordably programmed to cleave specific target sequences by binding short guide RNAs (sgRNAs) (Cong et al., 2013; Ran et al., 2013; Hsu et al., 2014). Another major advantage of CRISPR/Cas9 mutagenesis over previous gene targeting methods is its speed and efficiency (Mali et al., 2013; Ran et al., 2013; Fei et al., 2014). Target site cutting occurs minutes or hours after the Cas9-sgRNA complex enters the cytoplasm. In the case of zygotes, this means that the Cas9-sgRNA complex acts before and during early cell cleavage stages, creating mosaic mutant individuals (Mizuno et al., 2014; Yen et al., 2014). With highly efficient sgRNAs, most of the cells in these "F0" mosaic mutant individuals (sometimes called "crispants" (Burger et al., 2016)) will possess biallelic deletions in the targeted sequence (e.g. Blitz et al., 2013; Jao et al., 2013). CRISPR/Cas9 mutagenesis is also extremely versatile (Mali et al., 2013; Hsu et al., 2014). Despite evolving as a component of the prokaryotic adaptive immune system (Barrangou et al., 2007; Deveau et al., 2010; Garneau et al., 2010; Horvath and Barrangou, 2010; Makarova et al., 2011), Cas9 endonuclease appears to function efficiently and specifically in any cell type regardless of species. Thus, CRISPR/Cas9 allows the rapid creation of near-null mutants in virtually any organism whose eggs and/or zygotes are amenable to injection with proteins and RNA.

The efficiency and versatility of CRISPR/Cas9 means that genotype and phenotype can now be linked in organisms not suitable for the establishment of purebred lines. Thus, gene function can now be studied in organisms chosen for features aside from their ability to be lab-reared, including phylogenetic position, possession of derived phenotypes, or similarity to ancestral forms (Stolfi et al., 2014; Square et al., 2015; Trible et al., 2017; Rasys et al., 2019; Crawford et al., 2020; Kiyonari et al., 2021; Mori and Nakamura, 2021). This has opened up the possibility of side-by-side comparisons of gene function across diverse taxa, allowing researchers to more easily deduce ancestral gene functions, and identify the genes underlying novel phenotypes. CRISPR/Cas9 mutagenesis has become a powerful tool for understanding the evolution of genes, genomes, and phenotypes across both large and small evolutionary timescales (Komor et al., 2017).

CRISPR/Cas9 mutagenesis is having a large impact on the study of vertebrate evolution (e.g. Barske et al., 2020; Square et al., 2020; Hawkins et al., 2021). This is because the majority of vertebrates are large, long-lived, slow-maturing organisms, and do not survive well in small laboratory enclosures. This is especially true for taxa that diverged from the two dominant modern lineages, teleosts and tetrapods, in the Ordovician, or before. These so-called "basal" fish include the living jawless fish, hagfish and lamprey, and non-teleost jawed fish such as sturgeon, paddlefish, gar, bichir and bowfin. Unlike zebrafish and mouse, these vertebrates are typically large, with long generation times and extended predatory adulthoods.

Extant sturgeons are the few remaining representatives of once-diverse radiation of non-teleost ray-finned fish, the order acipenseriforms, that diverged from the lineage leading to modern teleosts about 345 million years ago (Hughes et al., 2018; Du et al., 2020). Because of their phylogenetic position, comparisons between sturgeons and teleost models, like zebrafish and medaka, can provide insights into the biology of early ray-finned fish (e.g. Minarik et al., 2017; Stundl et al., 2020). In addition, Acipenseriforms possess a combination of ancestral and derived vertebrate traits not seen in teleosts, including an endoskeleton lacking proper bone, a body armor made of bony scutes, a heterocercal caudal fin, and a lack of teeth in adulthood (Bemis et al., 1997). Several species are also polyploid (Havelka et al., 2013; Rajkov et al., 2014; Symonová et al., 2017) and roe of some is considered a delicacy (Bemis et al., 1997). Together, these characters make sturgeons an object of study for scientists from diverse specializations, from genetics and genomics to developmental and evolutionary biology, to aquaculture and food production. Despite this broad interest, understanding the genetic bases of sturgeon traits is difficult because they are large, long-lived, slow-maturing, and poorly suited for the establishment of purebred lines.

We recently adapted the CRISPR/Cas9 method to the sea lamprey, *Petromyzon marinus*, and the African clawed frog, *Xenopus laevis* (Square et al., 2015; Square et al., 2020), to better understand the ancestral functions of vertebrate developmental regulatory genes. Here we report the successful application of the same strategy to the non-teleost jawed fish sterlet, *Acipenser ruthenus,* a small species of sturgeon (**Figure 1A**). We then discuss specific variables and general considerations for workers seeking to apply the CRISPR/Cas9 mutagenesis method to other non-teleost fish, or any aquatic vertebrate for which establishing purebred lines is difficult or impossible.

MATERIALS AND METHODS

Animal Husbandry and *in vitro* Fertilization

We obtained the zygotes of sterlet (*Acipenser ruthenus* Linnaeus, 1758) from the adults kept and regularly bred at the Research Institute of Fish Culture and Hydrobiology, Faculty of Fisheries and Protection of Waters, University of South Bohemia in České Budějovice, Vodňany, Czech Republic (RIFCH). The husbandry, animal conditioning, gamete collection, and fertilization were described in detail by Chebanov and Galich (2011) and Saito et al. (2014). The adult breeding fish were handled under anesthesia in 0.05% tricaine. Briefly, both female and male adult sterlets, aged five to 9 years, were transferred from the outdoor ponds into 4,000 L indoor tanks with water temperature kept at constant 15°C. Spermiation in males was induced by intramuscular injection of carp pituitary extract at 4 mg/kg body weight in 0.9% NaCl solution. Sperm was collected 48 h later via a 0.6 mm catheter into cell culture flasks and kept on ice until *in vitro* fertilization. Female ovulation was induced in a similar way, but with two doses of the hormone instead of only one injected 12 h apart −0.5 and 4.5 mg/kg of body weight, respectively. Females ovulated 18–20 h after the hormone was administered. Ovulating females were placed in dorsal recumbency position and oviduct incision was performed using an eye microsurgery scalpel. This minimal invasive procedure allows accessing the ovulated eggs in sturgeon oviduct that is physiologically folded (see Chebanov and Galich, 2011). Then the female's abdomen was massaged in anterior-to-posterior direction and eggs were collected into bowls (**Figures 1B,C**), sealed with aluminum foil and subsequently fertilized with sperm at 15°C in dechlorinated tap water. We only used sperm with the spermatozoa motility assessed at >80% (Chebanov and Galich, 2011). The zygotes were rinsed in 0.04% tannic acid to make them less adherent (Saito et al., 2014), and then kept in water at 15°C before they were used in experiments. After the injections, we transferred the embryos to 48-well plates, each in 1 ml volume of E2 zebrafish medium (Brand et al., 2002) containing antibiotics (120 ng/ml of penicillin and 200 ng/ml of streptomycin) kept at 15°C. The medium was changed daily or more often, if necessary. Embryos selected for raising to later developmental stages were transferred to well-oxygenated tanks with E2/antibiotics at 15–17°C shortly after they hatched, which typically happened 7 days post fertilization (st. 35). We staged the embryos and larvae using the staging system of Detlaff et al. (1993). When the embryos and larvae reached the desired developmental stage, they were anesthetized by tricaine (MS-222; Serva) and fixed in 4% PFA in PBS overnight at 4°C. After several washes in PBS, we gradually dehydrated the embryos and larvae through a series of PBS/methanol solutions and stored in 100% methanol at −20°C until further use.

Animal care and all experiments were approved by the Ministry of Agriculture of the Czech Republic (MSMT-12550/2016-3), followed the principles of the European Union Harmonized Animal Welfare Act of the Czech Republic, and Principles of Laboratory Animal Care and National Laws 246/1992 "Animal Welfare", and were conducted in accordance with the Animal Research Committee of RIFCH. Authors of the study own the Certificate of professional competence for designing experiments and experimental projects under Section 15 d (3) of the Czech Republic Act no. 246/1992 Coll. on the Protection of Animals against Cruelty.

Identification of Endogenous *Tyr* and *Shh* Loci in Sterlet Genome and Design of the sgRNAs

We used similar F0 mutagenesis strategy as was described for sea lamprey (*Petromyzon marinus*) and African clawed frog (*Xenopus laevis*) (Square et al., 2015; Square et al., 2020) (**Figure 1D**). Using the spotted gar (*Lepisosteus ocelatus*) and zebrafish (*Danio rerio*) tyrosinase and sonic hedgehog protein sequences as queries we searched sterlet *Tyr* and *Shh* homologs in our *de novo* assembled transcriptomes of sterlet pharyngulae (available online at https://www.researchgate.net/profile/David-Jandzik/projects). The identity of recovered sequences was checked with BLAST (Altschul et al., 1990). We targeted the protein-coding exons at loci showing high evolutionary conservation across vertebrate taxa. In both *tyr* and *shh* we identified four putative exons. We selected the best CRISPR/Cas9 target sites with sequence 5′-GG (18N) NGG-3′ and with no off-target matches to our transcriptomes. The individual sequences were visually checked and identified as off-targets if they showed more than 85% similarity to our candidate sequence including PAM by BLAST (0–3 mismatches) and the mismatches were close to PAM site (more than one mismatch closer to PAM than 10 bp). The sequences of target sites were as follows: *tyr* sgRNA 3: GGTTAGAGACTTTATGTAAC (GGG), *tyr* sgRNA 4: GGCTCCATGTCTCAAGTCCA(AGG), *shh* gRNA1: (CCC) CAATGTGGCCGAGAAGACCC, *shh* gRNA2: GGGCCAGTG GCAGATATGAA(GGG) with PAM sites in parentheses and the PCR primers used to amplify the target sequences in **Table 1**. The amplified fragments were annealed and *in vitro* phosphorylated with T4 Polynuclease Kinase (NEB M0201S) at 37°C for 1h, and ligated into the DR274 plasmid (Hwang et al., 2013) pre-digested with BsaI (NEB R0535). Single guide RNAs were *in vitro* transcribed using T7 High Yield Kit (New England Biolabs) and purified by phenol-chloroform extraction followed by precipitation in 70% ethanol with 0.3M sodium acetate. The precipitate was resuspended in nuclease-free water.

TABLE 1 | Primers used A) to amplify the DNA templates for sgRNAs synthesis, B) for genotyping the putative mutants, and C) to amplify the DNA template for *foxD3, ripply3*, and *twist1* RNA *in situ* hybridization probe.

Target	Forward primer	Reverse primer
A)		
tyr sgRNA 3	TAGGTTAGAGACTTTATGTAAC	AAACGTTACATAAAGTCTCTAA
tyr sgRNA 4	TAGGCTCCATGTCTCAAGTCCA	AAACTGGACTTGAGACATGGAG
shh sgRNA 1	TAGGGTCTTCTCGGCCACATTG	AAACCAATGTGGCCGAGAAGAC
shh sgRNA 2	TAGGGCCAGTGGCAGATATGAA	AAACTTCATATCTGCCACTGGC
B)		
tyr sgRNA 3 geno	GGAGGAAGCAAACAACATAAGCTACAG	CACGGATATGACTGGAGGTAACAGTC
tyr sgRNA 4 geno	GCAGTTTACTTTGCTGCATGTGTG	CCACGTGGCTGTCTATCGGTG
shh sgRNA 1and2 geno	CTTTGGTGTCCTCTGGGCTG	GAGCCTGTCAGCCCCAGTG
C)		
foxD3 ISH probe	GAYGTGGAYATCGAYGTGGT	CTSARRAARCTVCCGTTGTC
ripply3 ISH probe	AGATGCAATCCACGGGCTAC	GTGGATTGTCGCTTGCACAG
twist1 ISH probe	GAAAWGWTGCARGANGAATC	TGVGATGYRGACATGGACCA

Microinjection

We prepared a fresh injection mix on ice shortly before each injection session. The commercially produced recombinant Cas9 protein (PNA Bio Inc.) was resuspended per manufacturer's instructions to the stock concentration of 1 mg/ml, aliquoted, and stored at −80°C. For a 6 μl injection mix we first incubated 1.6 μg of diluted Cas9 with 800 ng of total sgRNA for 10 min on ice, then brought the total volume up to 5.5 μl with nuclease-free water. Approximately 0.6 μl of 50 μg/μl lysinated Rhodamine-dextran (LRD; Invitrogen) in nuclease-free water was then added, resulting in a final LRD concentration of 5 μg/μl. One-cell-stage embryos were manually dechorionated using Dumont forceps and positioned in shallow holes in modeling clay in a Petri dish to facilitate their proper orientation and stability. The microinjection was performed either with mouth or manual injector (set to 100 hPa for 1 s) using microcapillary needles (Drummond Microcaps) pulled in a Narishige pc-10 puller (58°C with two weight elements; diameter 1.02 mm). The needle tip diameter was adjusted to allow to produce a ~20 nl drop (approximately 1/7 of the sterlet zygote in diameter) containing around 2.67 ng of sgRNA and 5.3 ng Cas9 protein (in 1:2 wight ratio) in total. While injecting, we targeted the animal pole of the embryo at a ~45° angle (**Figure 1D**). After injection the embryos were transferred to 48-well plates with E2/antibiotics at 15°C. We screened the embryos for LRD at the neurula stage (ca. st. 21). To control for mortality rates, we injected in parallel a few batches of embryos with a non-functional sgRNA at the same mix composition and concentrations as the experimental mixes. Each sgRNA was injected multiple times, in eggs from multiple clutches and by several authors of this study.

Analysis of Phenotypes

We assessed the efficiency of *tyr* mutagenesis in injected larvae raised approximately to 16 mm of total length, when pigmentation is visibly present and conspicuous on the head and body of the wild-type individuals. Based on the severity of pigmentation reduction we scored the observed phenotypes in four categories as 0–25% reduction, 25–50% reduction, 50–75% reduction, and 75–100% reduction. We also checked the larvae for any non-specific malformations and deformities. The embryos injected with *shh* sgRNAs with Cas9 protein were fixed at the pharyngula stage (st. 28), when several organs regulated by hedgehog signaling have formed. These include heart, pre-oral gut, olfactory and optic placodes, first and second pharyngeal pouches, fore-, mid-, and hindbrain, somites and pronephros. Rather than scoring the global phenotypes of the embryos and comparing the severity of the phenotypes among different characters, we scored each structure separately, recording whether it was present or reduced/deformed. We also used *in situ* hybridization to visualize changes in expression patterns of neural crest markers *foxd3* and *twist1* and pharyngeal pouch formation marker *ripply3* in mutant embryos.

Genotyping

To confirm mutations of targeted loci, we genotyped selected sterlet embryos and larvae showing variable phenotypes; at least two individuals per sgRNA. Due to higher percentage of individuals showing no obvious reduction in pigmentation, i.e. wild-type-looking phenotype in *tyr* sgRNA injected larvae, we also genotyped a few of those to obtain a better picture of *tyr* sgRNA efficiency. In *tyr* sgRNA injected larvae we used tail clips obtained from freshly euthanized larvae, while in *shh* mutants we used the whole embryos after we photographed and analyzed their phenotypes or gene expression patterns. We digested the tissue with Proteinase K (80 IU/ml; Sigma-Aldrich) in 1X PCR buffer at 55°C for 12 h. The obtained genomic DNA extracts served as templates in PCR amplification reactions using GoTaq polymerase (Promega) with amplification primers listed in **Table 1**. The PCR program followed the manufacturer's recommendations with the primer-specific annealing temperatures of 57.5° for both loci. We subcloned the resulting amplicons into pJet1.2 plasmid (Thermo Fisher) and fragments obtained from purified colony PCR reactions were sequenced using M13 forward and reverse primers. The alignments of mutant and wild-type sequences were prepared manually. We considered an individual to bear a mutant genotype if at least two sequences represented mutant alleles (see **Table 4**).

In situ Hybridization

Whole mount *in situ* hybridization (ISH) was carried out as described in detail by Minarik et al. (2017). We retrieved the putative sterlet homolog of *ripply3* from our pharyngula transcriptomes. The 317-bp long DNA template sequence was PCR amplified from sterlet cDNA (amplification primers in **Table 1**) and subcloned into pGEM T-Easy vector (Promega) by standard procedures. Acquisition of *foxd3* and *twist1* sterlet sequences was described by Stundl et al. (2020). ISHs using injected and wild-type embryos were performed in separate tubes though in parallel and under the precisely same conditions to avoid variations in probe penetration and signal development.

Imaging

All photographs of sterlet embryos and larvae in PBS were taken with Olympus SZX12 stereoscopic microscope using z-stacking Deep Focus technology of QuickPhoto software (Promicra).

RESULTS

We used CRISPR/Cas9 system to induce insertions and deletions (indels) **Figure 1D** into protein-coding sequences of two sterlet genes; *tyr* and *shh*. *tyr* encodes the enzyme tyrosinase involved in melanin synthesis in vertebrates. Successful mutagenesis is expected to reduce pigmentation of sterlet larvae but should not affect other aspects of their normal development. On the other hand, *shh* is a developmental regulator of several organ systems, and its mutation is expected to have dramatic effects on early embryonic development of tissues and organ systems derived from neural crest cells and all three germ layers.

Tyr Mutagenesis

We designed and synthesized two sgRNAs targeting two different exons of the sterlet homolog of human *Tyr1* - sgRNA three and four against exons 2 and 3, respectively, sequences of which allowed us to design sgRNAs according to our criteria (see Methods and Square et al., 2015; Square et al., 2020). First, we co-injected 50 one-cell sterlet embryos with a mix containing both guides at a total amount of ~2.67 ng of sgRNA and 5.3 ng of Cas9 protein (in 1:2 wight ratio) per single embryo. The total amount was calculated to match the amount of guide RNA and Cas9 protein relative to the egg size successfully used in *Xenopus laevis* CRISPR/Cas9 mutagenesis (Square et al., 2015; Square et al., 2020). In total, 32 injected embryos were LRD positive (i.e., with the lysinated Rhodamine-dextran of the injection mix glowing under the fluorescent light) at the neurula stage, however, no larva showed reduction in pigmentation or other developmental malformation. We suspected that this could have resulted from concentration of each individual sgRNA being too low in this combined sgRNA mix and therefore we used the same total amount of the injection mix but with each sgRNA separately in the next experiment. Both these experiments produced albinos with the variable extent of pigment reduction and normal morphology lacking any visible morphological abnormalities. We visually assessed the extent of pigment reduction and scored the phenotypes to four classes at increments of 25% pigment reduction (**Table 2**). Injection with sgRNA three resulted in >70% partial or total albinism occurrence, while treatment with sgRNA four produced >50% partially or completely albinotic larvae. Mutagenesis with sgRNA three was more efficient in producing phenotypes with stronger pigment reduction, although the numbers of complete or near-complete albinos were similar between the guides. We recorded 25–33% mortality, which was similar to the uninjected dechorionated wild-type larvae raised to the same stage (**Table 2**) and to the mortality of embryos injected with non-functional sgRNA (14/50 = 28%).

TABLE 2 | Summary of observed phenotypes of Δ*tyr* sterlet larvae.

	***tyr* sgRNA 3**	***tyr* sgRNA 4**
0–25% pigment reduction	29% (6/21)	50% (9/18)
25–50% pigment reduction	29% (6/21)	26% (5/18)
50–75% pigment reduction	33% (7/21)	11% (2/18)
75–100% pigment reduction	10% (2/21)	11% (2/18)
Injected in total	40	30
LRD positive embryos	21	18
Mortality	10 (25%)	10 (33%)

We verified the disruption of the target *tyr* loci by genotyping selected individuals from each phenotype class (**Figure 2**, **Supplementary Figure S1**). We found mutant alleles in 7/10 and 4/7 genotyped larvae generated with sgRNA 3 and 4, respectively, and confirmed mutants in all phenotype classes. Approximately half of all obtained sequences were mutant. We observed 27 unique mutant alleles (out of 36 sequences), 22 of which had only deletions, one had only insertion and four had both deletions and insertions (**Supplementary Figure S1**). While we did not observe a correlation between the relative occurrence of WT vs. mutant alleles and severity of the observed phenotypes, it appears that the stronger phenotypes have more alleles with indels that cause frameshifts (**Figure 2**, **Supplementary Figure S1**).

Shh Mutagenesis

Similar to our initial experiments with *tyr* mutagenesis, we first injected 50 single-cell sterlet embryos with a mix of both guides targeting *shh* locus with the total amount of ~2.67 ng of sgRNA and 5.3 ng of Cas9 protein per embryo. We recorded relatively high mortality (36/50 = 72%), but the surviving LRD positive embryos (13; one was LRD negative) showed no discernable phenotypic effect. Next, we injected each sgRNA separately at the same total amount and sgRNA/Cas9 ratio. Due to high mortality in the initial experiment, we injected higher numbers with each guide in a separate mix - 191 with sgRNA 1 and 117 with sg RNA 2. We observed mortality of 34 and 38%, respectively (64/191 individuals in sgRNA one and 44/117 in sgRNA two injections) and obtained 93 and 57 LRD positive embryos that reached developmental st. 28. The embryos were severely affected with strong morphological deformations in almost all embryonic structures present at that stage (**Table 3**). The highest frequency phenotypes included missing pharyngeal pouches and pre-oral gut (up to 81%), followed by missing, reduced or

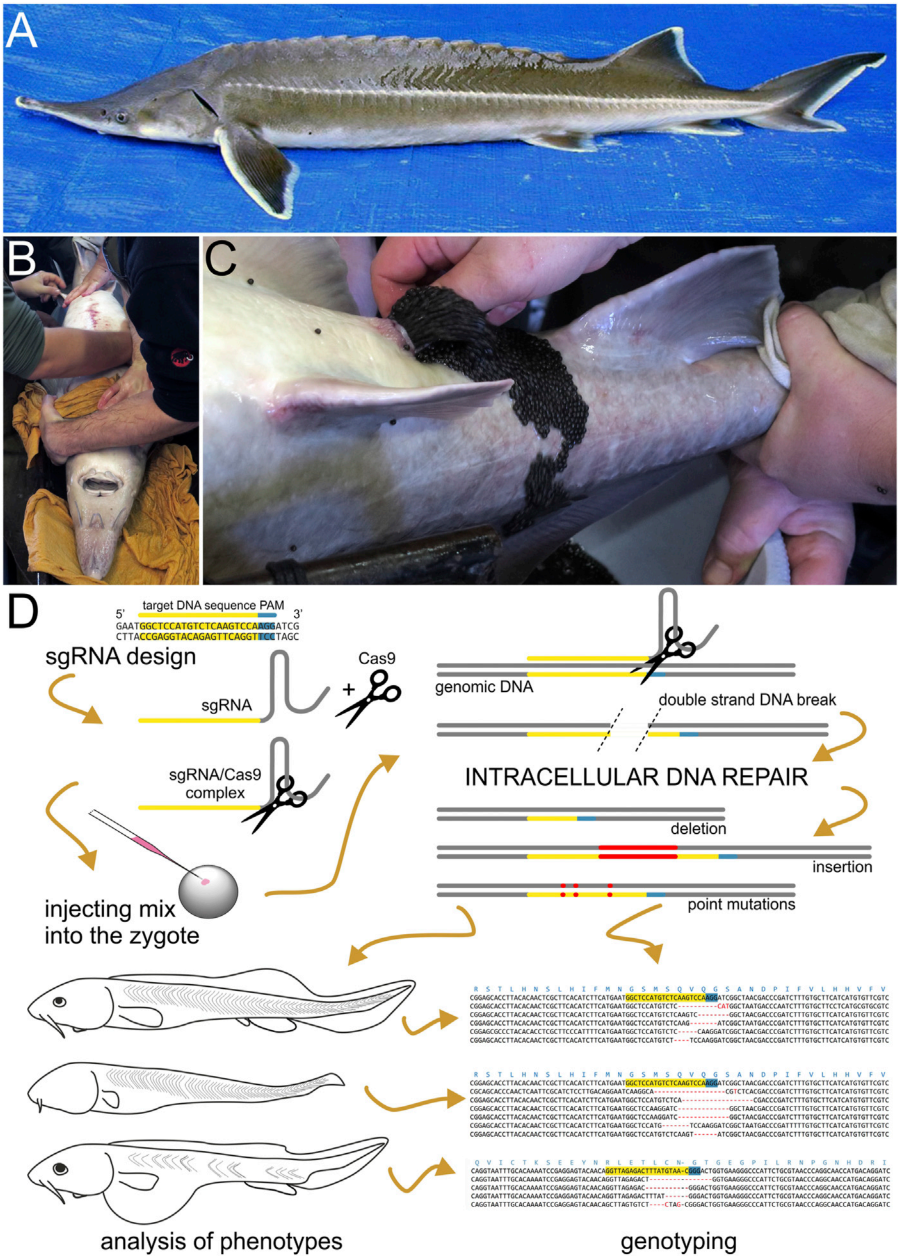

A
B
C
D
target DNA sequence PAM
5'
3'
GAATGGCTCCATGTCTCAAGTCCAAGGATCG
CTTACCGAGGTACAGAGTTCAGGTTCCTAGC
sgRNA design
sgRNA
Cas9
+
genomic DNA
double strand DNA break
sgRNA/Cas9 complex
INTRACELLULAR DNA REPAIR
deletion
insertion
point mutations
injecting mix into the zygote
analysis of phenotypes
genotyping

FIGURE 1 | Sturgeon and CRISPR/Cas9 mutagenesis pipeline. **(A)** Dorsoleteral view of an adult individual of sterlet (*Acipenser ruthenus*). Abdomen of an adult anesthetized sturgeon female (pictured here is the Siberian sturgeon, *Acipenser baerii*) in dorsal recumbency **(B)** is manually massaged in anterior-to-posterior direction and ovulated eggs are collected into a bowl **(C)**. **(D)** A scheme illustrating the sterlet CRISPR/Cas9 mutagenesis pipeline. First, the single guide RNA (sgRNA) is designed and synthesized based on the selected target DNA sequences and next to Protospacer Adjacent Motif (PAM). PAM is required for proper function of Cas9 nuclease and is not part of sgRNA. Next, the sgRNA is mixed with Cas9 protein (illustrated here as scissors) to form a sgRNA/Cas9 complex. This mix is injected into sterlet zygotes. Cas9 nuclease navigated to the target by sgRNA breaks the double strand of DNA, which is subsequently repaired inside the cleaving cells and nuclei. (mainly by non-homologous end joining) The imperfect DNA repairs introduce mutations into newly synthesized molecules of DNA resulting into a mosaic of cells bearing DNA with indels and/or point mutations at the targeted site. Embryos and/or raised larvae are later inspected for phenotypic effects and selected candidates are genotyped.

malformed optic placode, not discernible heart, pronephros reduction and brain malformations. Such deformations were consistent with patterns of *shh* gene expression observed during the sterlet embryonic stages preceding the analyzed stages (**Figures 3A–C**). Only 7–10% of injected LRD positive individuals appeared to have no visible phenotype (9/93 and 4/57 for sgRNA 1 and 2, respectively). We did not observe any of the described phenotypes in the control embryos and their mortality was similar to the mortality in *tyr* mutagenesis experiment.

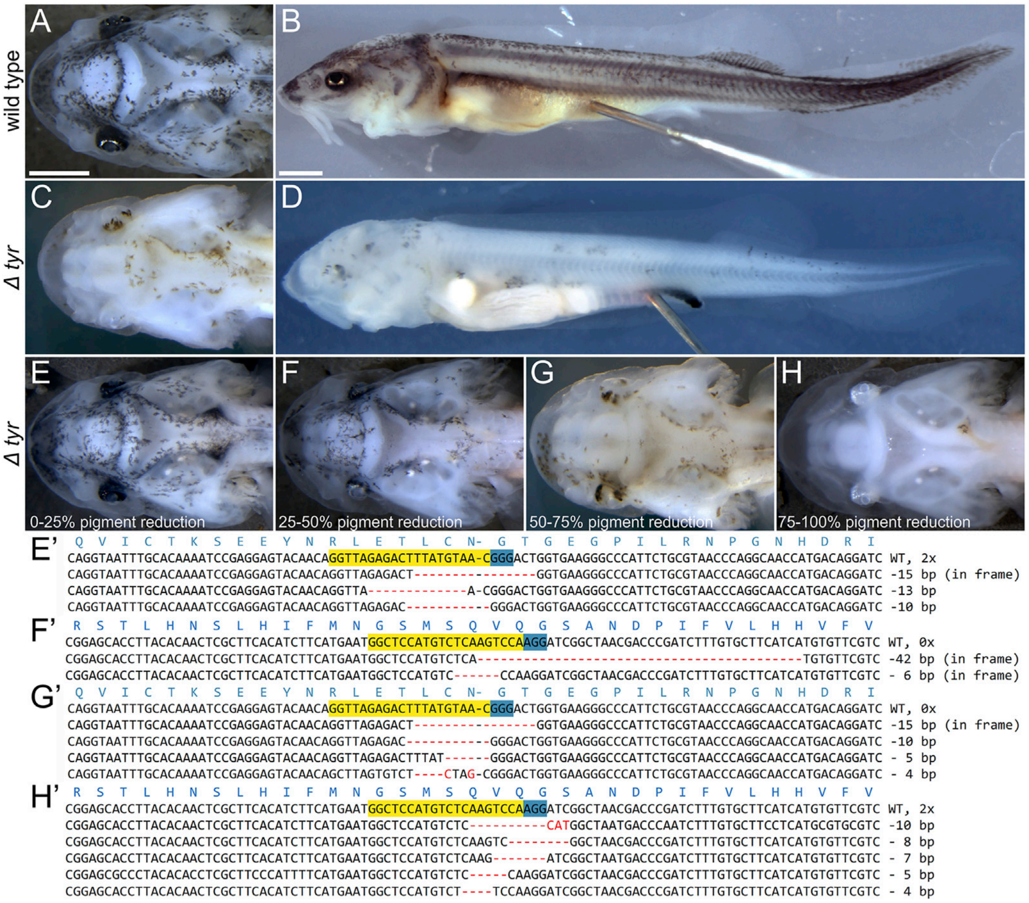

A
B
C
D
E
F
G
H
wild type
Δ tyr
Δ tyr
0-25% pigment reduction
25-50% pigment reduction
50-75% pigment reduction
75-100% pigment reduction
E’
Q V I C T K S E E Y N R L E T L C N- G T G E G P I L R N P G N H D R I
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGACTTTATGTAA-CGGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC WT, 2x
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGACT----------------GGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC -15 bp (in frame)
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTA------------A-CGGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC -13 bp
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGAC-----------GGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC -10 bp
F’
R S T L H N S L H I F M N G S M S Q V Q G S A N D P I F V L H H V F V
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTCAAGTCCAAGGATCGGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC WT, 0x
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTCA--TGTGTTCGTC -42 bp (in frame)
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTC------CCAAGGATCGGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC - 6 bp (in frame)
G’
Q V I C T K S E E Y N R L E T L C N- G T G E G P I L R N P G N H D R I
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGACTTTATGTAA-CGGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC WT, 0x
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGACT----------------GGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC -15 bp (in frame)
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGAC-----------GGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC -10 bp
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGGTTAGAGACTTTAT------GGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC - 5 bp
CAGGTAATTTGCACAAAATCCGAGGAGTACAACAGCTTAGTGTCT----CTAG-CGGGACTGGTGAAGGGCCCATTCTGCGTAACCCAGGCAACCATGACAGGATC - 4 bp
H’
R S T L H N S L H I F M N G S M S Q V Q G S A N D P I F V L H H V F V
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTCAAGTCCAAGGATCGGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC WT, 2x
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTC----------CATGGCTAATGACCCAATCTTTGTGCTTCCTCATGCGTGCGTC -10 bp
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTCAAGTC--------GGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC - 8 bp
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCTCAAG-------ATCGGCTAATGACCCGATCTTTGTGCTTCATCATGTGTTCGTC - 7 bp
CGGAGCGCCCTACACACCTCGCTTCCCATTTTCATGAATGGCTCCATGTCTC-----CAAGGATCGGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC - 5 bp
CGGAGCACCTTACACAACTCGCTTCACATCTTCATGAATGGCTCCATGTCT----TCCAAGGATCGGCTAACGACCCGATCTTTGTGCTTCATCATGTGTTCGTC - 4 bp

FIGURE 2 | Phenotypes and genotypes of Δ*tyr* sterlet larvae. Dorsal **(A,C,E–H)** and lateral **(B–D)** views of ~16 mm long larvae with anterior to the left. Wild type [WT; **(A,B)**] and representative mutant **(C,D)** larvae and four mutant classes **(E–H)** and their genotypes **(E′–H′)**. The CRISPR target sites and PAM on the forward sequences are highlighted in yellow and blue, respectively, and the red dashes and letters represent deletions, insertions, and polymorphisms relative to the WT sequence. Mutations in **(E′,G′)**, and **(F′,H′)** were introduced by *tyr* sgRNA three and sgRNA 4, respectively. Mutant and WT sequences were identified in all phenotype classes (see also **Supplementary Figure S1**), while sequencing of larvae portrayed on panels **(F,H)** only returned mutant sequences (as was the case of 3/11 mutant larvae). Scale bars (same in A + C + E-H and B + D) represent 1 mm.

Next, we analyzed effects of *shh* disruption on expression patterns in injected sterlet embryos by ISH. We examined three genes known to be involved in development of some of the affected structures at early embryonic stages: *foxd3* marking NCCs at early migration stage, *twist1* NCCs at later stages of migration, and *ripply3* in forming pharyngeal pouches (**Figure 4**, **Supplementary Figure S3**). We found that all these markers were reduced in the injected sterlets at st. 28. *foxd3* showed a reduction in the size of all three streams of migrating NCC (trigeminal, hyoid, and branchial), *twist1* expression was reduced in all later NCC populations (maxillary, mandibular and branchial) and even entirely missing in some embryos. Only three out of 14 hybridized embryos appeared to have a wild-type expression pattern of the NCC markers. *ripply3* ISH confirmed our observation of the strong effect of *shh* mutation on pharyngeal pouch formation in putative *shh* mutants; all seven analyzed embryos had either missing or reduced expression domains (**Table 3**).

To verify mutagenesis in *shh* locus in analyzed sterlet embryos, we genotyped selected individuals injected by both sgRNAs, including those used for ISHs. All 33 obtained sequences had indels and consequently all nine genotyped individuals were confirmed mutants. We identified 21 unique alleles in total, five of which only showed insertions, 12 of them contained deletions and four had both insertions and deletions (**Table 4**). The majority of the observed mutations caused frameshifts; these were observed in at least half of all obtained sequences of each mutant individual. Only 8/26 sequences of mutations introduced by sgRNA one were in-frame (**Figure 4**, **Figure 3**, **Supplementary Figure S2**, **Table 4**).

DISCUSSION

Sterlet is Highly Amenable to CRISPR/Cas9 F0 Mutagenesis

Here we report the application of a CRISPR/Cas9 mutagenesis strategy previously optimized in the sea lamprey (Square et al.,

TABLE 3 | Summary of observed phenotypic effects of Δ*shh* sterlet embryos.

A)	***shh* sgRNA 1**		***shh* sgRNA 2**	
	Missing/not visible	**Reduced/malformed**	**Missing**	**Reduced/malformed**
Heart	23% (21/93)	—	16% (9/57)	—
Pre-oral gut	47% (44/93)	—	46% (26/57)	—
Olfactory placode	13% (12/93)	8% (7/93)	12% (7/57)	9% (5/57)
Optic placode	22% (20/93)	34% (32/93)	23% (13/57)	30% (23/57)
Phar. pouch 1	73% (68/93)	—	81% (46/57)	—
Phar. pouch 2	69% (64/93)	—	68% (39/57)	—
Forebrain	9% (8/93)	24% (22/93)	7% (4/57)	21% (12/57)
Midbrain	6% (6/93)	17% (16/93)	7% (4/57)	12% (7/57)
Hindbrain	6% (6/93)	18% (17/93)	4% (2/57)	14% (8/57)
Somites	9% (8/93)	—	5% (3/57)	—
Pronefros	—	33% (31/93)	—	26% (15/57)
100% Healthy embryos	10% (9/93)		7% (4/57)	

B)	**missing**	***Shh* mutant**	**reduced**
foxd3 ISH			
-Trigeminal NC	0/3		3/3
-Hyoid NC	0/3		3/3
-Branchial NC	0/3		3/3
100% Healthy embryos		0/3	
twist1 ISH			
-Maxillary NC	1/14		3/14
-Mandibular NC	4/14		3/14
-Hyoid NC	1/14		9/14
-Branchial 1 NC	1/14		7/14
-Branchial 2 NC	5/14		5/14
100% Healthy embryos		3/14	
ripply3 ISH			
-Pharyngeal pouch 1	4/7		3/7
-Pharyngeal pouch 2	0/7		7/7
-Pharyngeal pouch 3	4/7		3/7
100% Healthy embryos		0/7	

2015) and African clawed frog (Square et al., 2020), to the sterlet. Previous reports using TALENs and CRISPR/Cas9 suggested sterlet zygotes will tolerate injection with proteins and nucleic acids and were amenable to gene targeting (Chen et al., 2018; Baloch et al., 2019). We found that our optimized protocol introduced biallelic mutations in the targeted loci at relatively high efficiency and yielded consistent and reproducible phenotypes. As expected, targeting of the gene *tyr*, resulted in larval albinism with no other discernible developmental defects. These mutants have similar mortality as their uninjected wild type siblings. The visual scoring of the larval phenotypes is straightforward, suggesting that *tyr* targeting could serve as a robust positive control in future studies. In contrast to *tyr* disruption, mutating the pleiotropic developmental regulator *shh* caused defects in a wide range of embryonic and larval tissues and structures. Interestingly, both sgRNAs used to target *shh* resulted in all injected individuals displaying defects consistent with the known roles of *shh* in vertebrate development. This demonstrates that determining the function of developmental regulators, even highly pleiotropic ones, is readily achievable in sterlet. Arguably, in CRISPR/Cas9 sterlet F0 mutants it could be more challenging to detect and score very subtle and less conspicuous phenotypes than in models with stable inbred lines.

Though sterlet is not well suited as a laboratory-propagated model organism, several features of its natural history make it a strong candidate for routine genetic analyses of "F0" mosaic mutants. Along with other sturgeon species, it is a frequently farmed fish, whose eggs can be seasonally obtained in large quantities. The eggs are relatively large and can be easily manipulated, dechorionated, and injected with just forceps and also simple hand-held and mouth operated injector. The embryos and larvae thrive in simple aquaria with no special care requirements beyond clean oxygenated water and a constant temperature of 15–17°C. Several common laboratory methods have also already been optimized in the species including histological sectioning, microCT scanning, RNA *in situ* hybridization, lineage tracing following injections, and immunohistochemistry (Minarik et al., 2017; Stundl et al., 2020).

Other features of sterlet, and acipenseriform fishes in general, were likely important for the successful application of our CRISPR/Cas9 mutagenesis strategy. Sterlet eggs are rich in yolk that is distributed throughout the egg (Dettlaff et al., 1993). The eggs of lamprey and frog, which are highly amenable to CRISPR/Cas9 mutagenesis (Square et al., 2015; Square et al., 2020), are similarly rich in yolk. Since as in frog, the sterlet egg yolk is slightly more concentrated on the vegetal pole, the animal pole with more yolk-free cytoplasm and nucleus

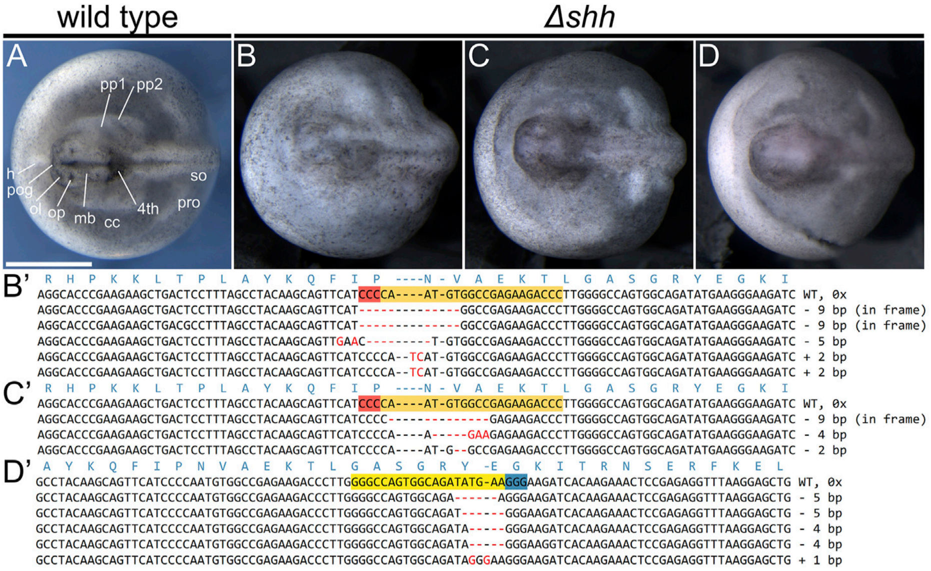

wild type
Δshh
A
B
C
D
pp1
pp2
h
pog
ol
op
mb
cc
4th
so
pro
B’
R H P K K L T P L A Y K Q F I P ----N - V A E K T L G A S G R Y E G K I
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC WT, 0x
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCAT--------------GGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 9 bp (in frame)
AGGCACCCGAAGAAGCTGACGCCTTTAGCCTACAAGCAGTTCAT--------------GGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 9 bp (in frame)
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTGAAC---------T-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 5 bp
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA--TCAT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 2 bp
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA--TCAT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 2 bp
C’
R H P K K L T P L A Y K Q F I P ----N - V A E K T L G A S G R Y E G K I
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC WT, 0x
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCC--------------GAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 9 bp (in frame)
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----A-----GAAGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 4 bp
AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-G--GCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 2 bp
D’
A Y K Q F I P N V A E K T L G A S G R Y -E G K I T R N S E R F K E L
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATG-AAGGGAAGATCACAAGAAACTCCGAGAGGTTTAAGGAGCTG WT, 0x
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGA------AGGGAAGATCACAAGAAACTCCGAGAGGTTTAAGGAGCTG - 5 bp
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGAT------GGGAAGATCACAAGAAACTCCGAGAGGTTTAAGGAGCTG - 5 bp
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATA-----GGGAAGATCACAAGAAACTCCGAGAGGTTTAAGGAGCTG - 4 bp
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATA-----GGGAAGGTCACAAGAAACTCCGAGAGGTTTAAGGAGCTG - 4 bp
GCCTACAAGCAGTTCATCCCCAATGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATAGGGAAGGGAAGATCACAAGAAACTCCGAGAGGTTTAAGGAGCTG + 1 bp

FIGURE 3 | Phenotypes and genotypes of Δ*shh* sterlet larvae. Dorsal views of st. 28 larvae with anterior to the left. Wild type [WT; **(A)**] and mutant **(B–D)** embryos and their mutant alleles **(B′–D′)**. The CRISPR target sites are highlighted in orange with red PAM on the reverse strand sequences, while they are yellow with blue PAM on the forward strand. The red dashes and letters represent deletions, insertions, and polymorphisms relative to the WT sequence. Mutations in **(B′,C′)** were introduced using *shh* sgRNA 1, while the mutations in **(D′)** resulted from injections of *shh* sgRNA 2. No WT sequences were identified in any of the individuals showing *shh* phenotype (six genotyped individuals in total). The mutant embryo in **(B)** was scored as missing pre-oral gut, optic and olfactory placodes, forebrain, midbrain, part of the hindbrain, and pharyngeal pouches and with underdeveloped pronephros. The mutant embryo in **(C)** is missing pharyngeal pouches and has deformed forebrain and midbrain. The mutant embryo in **(D)** was identified as missing pharyngeal pouches and somites, the entire head is underdeveloped. Fourth—fourth brain chamber of hindbrain, cc—coelomic cavity, h—heart, mb—midbrain, ol—olfactory placode, op—optic placode, pog—pre-oral gut, pp—pharyngeal pouch, pro—pronephros, so—somites. Scale bar represents 1 mm.

are better accessible for the injected CRISPR/Cas9 mix. Therefore, we tried to target the animal pole, usually conveniently situated on the top of the embryo. However, we suspect that because acipenseriforms undergo holoblastic cleavage (Dettlaff et al., 1993; Ostaszewska and Dabrowski, 2009) injecting any part of the zygote’s cytoplasm would likely work well. Similar to frog, but unlike lamprey, sturgeon eggs have relatively thick envelope. While this does not prevent injection of the embryo with a capillary needle, we preferred to remove the thick outer layer with forceps to better see the injected droplet size and target the animal pole (frog eggs jelly membranes are removed by cysteine treatment; Sive et al., 2000). Due to the similarities between the zygotes of sterlet and those of lamprey with frog, we decided to use approximately the same concentration of injected RNA and protein, and approximately the same relative volume of injection solution. In practical terms, this means we injected a droplet of CRISPR/Cas9 injection mixture with a diameter approximately 1/7 that of the zygote diameter. Because this strategy works well in sterlet, lamprey, and African clawed frog, we suspect a 1/7 size droplet would yield good results in any vertebrate with isolecithal, mesolecithal, or polylecithal eggs displaying holoblastic cleavage. In contrast, zygotes of embryos with meroblastic cleavage (e.g., Takeuchi et al., 2009), such as zebrafish and other teleost fish, appear to tolerate lower amounts of RNA/protein (Hwang et al., 2013; Jao et al., 2013), and lower volumes of injection mix. For such embryos, the injected volume should likely be adjusted based on the size of the blastodisc rather than the entire embryo.

Recent publication of high-quality sterlet genome (Du et al., 2020) should allow the effective design of highly specific sgRNAs. We initially only searched our embryonic transcriptomes for potential off-target sequences. However, additional searches of the new sterlet genome, including non-coding sequences, did not identify any other off-target sequences. When designing sgRNAs, we found that approximately one third of the initially selected guide sequences conformed with off-target sequences according to the criteria described in Methods. When performing CRISPR/Cas9 mutagenesis in sea lamprey and African clawed frog we found that using at least two different sgRNAs per gene, and injecting them separately, is the best strategy for producing verifiable and consistent phenotypes and control for any defects caused by off-target lesions. We observed highly consistent phenotypes and no indication of off-target effects in sea lamprey after mutating more than 20 different genes (Square et al., 2020). The results presented here suggest a similar level of specificity and consistently can be expected in sterlet.

We only observed a moderate disparity in the efficiency of individual sgRNAs. While the mutation rates varied between *tyr* and *shh* loci, they were very similar between the guides targeting the same gene (**Table 4**), with subtle differences recorded among mutant categories. This contrasts with sea lamprey and other species, in which some degree of variation was observed among different guides targeting the same genes (Hsu et al., 2013; Square et al., 2015). We expect that injecting more eggs and guides would result in higher variation in sgRNA efficiency in sterlet as well. On

TABLE 4 | Summary of sterlet Δ*tyr* larvae and Δ*shh* embryos genotyping. A) all obtained sequences from all genotyped individuals pooled together, B) unique alleles from all genotyped individuals pooled together, C) all genotyped individuals. The denominator numbers in A represent all obtained sequences from all individuals genotyped for a respective locus, and similarly in B, the denominators show the total number of mutant alleles obtained from sequencing all individuals. In C, the denominators indicate the total numbers of genotyped individuals for each respective locus. Mutation and frameshift rates are calculated for each row, i.e. they represent percentages of mutant sequences and unique alleles from all sequences obtained by genotyping all individuals pooled together in A and B, respectively, while they show percentages of genotyped individuals with mutant/frameshifted genotypes in C.

	Mutant	WT	Frameshift	Mutation rate	Frameshift rate (%)
A) Sequences					
tyr sgRNA 3	21/46	25/46	11/21	46%	52
tyr sgRNA 4	15/30	15/30	8/15	50%	53
shh sgRNA 1	26/26	0/26	18/26	100%	69
shh sgRNA 2	7/7	0/7	7/7	100%	100
B) Unique alleles					
tyr sgRNA 3	14/21	—	7/14	—	50
tyr sgRNA 4	13/15	—	8/13	—	62
shh sgRNA 1	15/26		11/15		73
shh sgRNA 2	6/7	—	6/6	—	100
C) Individuals					
tyr sgRNA 3	7/10	3/10	6/7	70%	86
tyr sgRNA 4	4/7	3/7	2/4	57%	50
shh sgRNA 1	7/7	0/7	7/7	100%	100
shh sgRNA 2	2/2	0/2	2/2	100%	100

the other hand, as we observed in sea lamprey, mutations causing frameshifts in coding DNA sequences correlated with more severe phenotypes in sterlet as well.

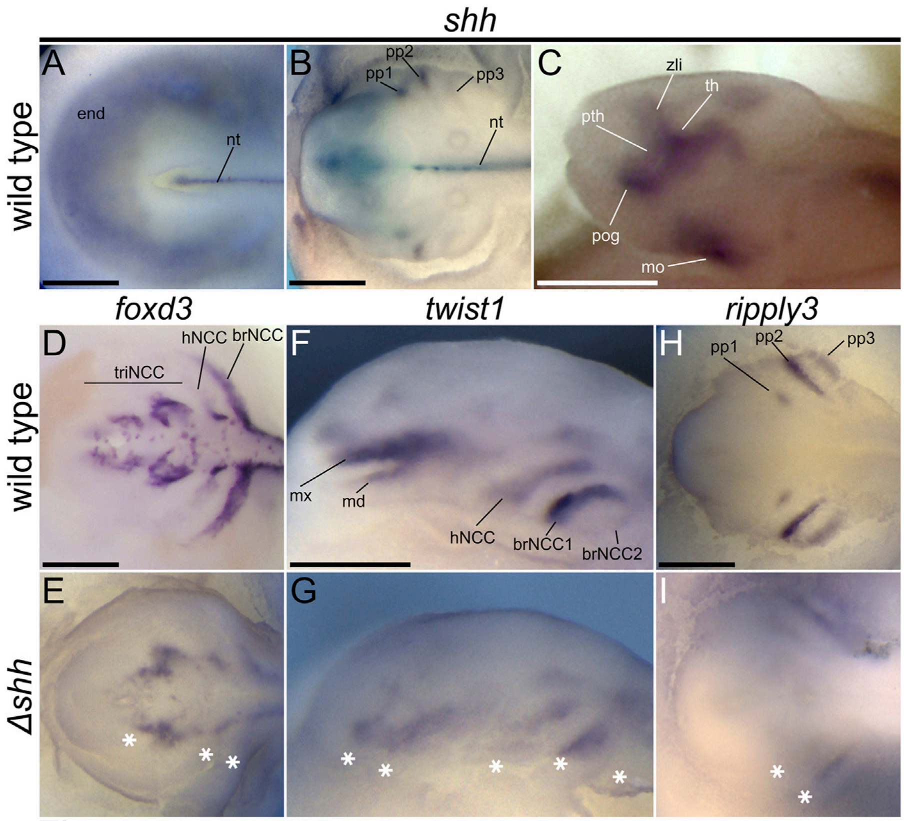
To confirm successful mutagenesis, we also genotyped representative individuals from each phenotypic class. While PCR fragment length is often used to tell whether an individual harbors mutant alleles, we find that this method often produces ambiguous results. We thus chose to clone and sequence genomic fragments including the target sequence and approximately ~200 bp of flanking sequence on each side. Sequencing of phenotypic mutants and controls, confirmed highly efficient mutagenesis with all tested sgRNAs, with the frequency and type of alleles correlating with phenotype. While even two to three of sequences per individual are sufficient to confirm a baseline level of mutagenesis, we find sequencing ten or more clones gives a view of allelic composition that usually correlates with phenotypes, and also suggests the efficiency of the sgRNA. We also found that, as with lamprey embryos (York et al., 2017), fixed *in situ* hybridized sterlet embryos can be successfully genotyped if they are not post-fixed with formaldehyde for extended periods of time (see also Square et al., 2020).

Prospects for Adapting CRISPR/Cas9 F0 Mutagenesis to Other "Basal" Fishes

The efficiency, precision, and ease of use of CRISPR/Cas9 mutagenesis in sterlet, lamprey, and African clawed frog strongly suggest this method can be successfully applied to other non-teleost fish. Besides sterlet, several other sturgeon species are currently farmed. While sterlet is probably the most suitable model for developmental studies due to its smaller size and monoploid genome, several other species with different levels of polyploidy (Havelka et al., 2016) could be attractive for studies of genome evolution. While polyploidy necessarily complicates targeted gene mutagenesis by requiring simultaneous disruption of homeologs to yield a valid loss-of-function, the efficiency of CRISPR/Cas9 can be exploited to overcome this issue. We and others have found that targeting conserved homeolog sequences, and/or using multiple sgRNAs can yield strong phenotypes in the allotetraploid African clawed frog (Wang et al., 2015; Square et al., 2020). Gene manipulation in larger species of sturgeons can potentially be useful in sturgeon food production industry when targeting genes involved in growth, immunity, or egg production.

American paddlefish (*Polyodon spathula*) is morphologically unusual acipenseriform relative of sturgeons. Like sturgeon, it is also frequently farmed, and its embryos have been used in several evolutionary developmental studies (e.g. Davis et al., 2007; Modrell et al., 2011). There are currently no published reports of CRISPR/Cas9 mutagenesis in paddlefish. However, high similarity of paddlefish and sterlet embryos, and successful previous experiments involving dye injections at later embryonic stages suggest that CRISPR/Cas9 mutagenesis could be highly effective in paddlefish. Unfortunately, collecting of early paddlefish embryos appears to be more challenging than sterlet in part because of the shorter paddlefish spawning season (Modrell et al., 2017).

Besides acipenseriforms, three non-teleost fish lineages have extant members that have been utilized in comparative embryological studies and have published genomes - bichirs (Polypteriformes), gars (Lepisosteiformes), and bowfins (Amiiformes) (Askary et al., 2016; Braasch et al., 2016; Minarik et al., 2017; Stundl et al., 2019; Funk et al., 2020; Stundl et al., 2020; Thompson et al., 2021; Bi et al., 2021). While bichirs and bowfins are evolutionarily very attractive and informative species with unique morphologies, the main challenge in targeted mutagenesis in these



```
E’  R  H  P  K  K  L  T  P  L  A  Y  K  Q  F  I  P  ----N - V  A  E  K  T  L  G  A  S  G  R  Y  E  G  K  I
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC WT, 0x
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGT------------------GGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC -13 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGT------------------GGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC -13 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCGTCTT-CT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 3 bp (in frame)
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCGTCTT-CT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 3 bp (in frame)
G’  R  H  P  K  K  L  T  P  L  A  Y  K  Q  F  I  P  ----N - V  A  E  K  T  L  G  A  S  G  R  Y  E  G  K  I
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC WT, 0x
   AGGCGCCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCC---------------GAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC -10 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA--------------AGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 9 bp (in frame)
   AGGCACCCGAAGAAGCTGACTCCCTTAGCCTACAAGCAGTTCATCCCCA----A-----GAAGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 4 bp
I’  R  H  P  K  K  L  T  P  L  A  Y  K  Q  F  I  P  ----N - V  A  E  K  T  L  G  A  S  G  R  Y  E  G  K  I
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----AT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC WT, 0x
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATC-------------GGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 8 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCC-------------TCGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC - 8 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA--TCAT-GTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 2 bp
   AGGCACCCGAAGAAGCTGACTCCTTTAGCCTACAAGCAGTTCATCCCCA----ATTGTGGCCGAGAAGACCCTTGGGGCCAGTGGCAGATATGAAGGGAAGATC + 1 bp
```

FIGURE 4 | *shh* expression patterns and Δ*shh* sterlet embryos at st. 28 showing defects in expression patterns of *foxd3*, *twist1*, and *ripply3*. Dorsal **(A,B,D,E,H,I)** and lateral **(C,F,G)** views with anterior to the left. Patterns of *shh* expression in various morphological structures of the wild-type sterlet embryos at stages 23 **(A)**, 28 **(B)**, and 30 **(C)**. *foxd3* marks neural crest cells (NCCs) at early stages of migration, *twist1* marks NCC at later migration stages, and *ripply3* is expressed in developing pharyngeal pouches. Compare with **Supplementary Figure S3** to see the observed variation in expression patterns in both wild-type and mutant individuals. The sequences show the reverse DNA strands with *shh* sgRNA one CRISPR target sites highlighted in orange with red PAM site. The red dashes and letters show indels and polymorphisms relative to the WT sequence. Asterisks indicate missing or reduced expression in mutant embryos. brNCC—branchial stream of NCCs, hNCC—hyoid stream of NCCs, md—mandibular stream of trigeminal stream of NCCs, mo—mouth, mx—maxillary stream of trigeminal stream of NCCs, nt—notochord, pog—pre-oral gut, pth—prethalamus, pp—pharyngeal pouch, th—thalamus, triNCC—trigeminal stream of NCCs, zli—zona limitans intrathalamica. Scale bars represent 0.5 mm.

organisms is obtaining zygotes for injection. Bichirs reproduce in aquaculture settings and can occasionally spawn in home aquaria. However, their mating is secretive and usually occurs at night. Thus, collection of zygotes requires constant observation, which could disturb the animals and prevent spawning. Once collected, though, bichir embryos are relatively sturdy, and can be manipulated and injected with vital dyes much like African clawed frog (Stundl et al., 2019). Bowfin eggs are typically only collected in the wild, and their zygotes are thus less accessible than those of bichir (Funk et al., 2020). Fortunately, decapsulating of later bowfin embryonic stages is relatively easy, so it is likely that early embryos could be injected if obtained (Brent Hawkins, personal communication). Cleavage of bowfin embryos is intermediate between the holoblastic mode of most non-teleost fishes and the meroblastic cleavage of teleosts, with reduced cleavage on the vegetal hemisphere (Cooper and Virta, 2007). This would suggest that if bowfin zygotes were collected, a smaller injection volume might be needed for embryos to survive. Besides sturgeons, perhaps the most promising non-teleost fish candidates for CRISPR/Cas9 mutagenesis are gars. Several species are kept and propagated at fish farms, and like in sturgeons, large quantities of eggs can be seasonally obtained in a controlled fashion *in vitro* fertilization (Braasch et al., 2014 in spotted gar; Minarik et al., 2017 in tropical gar). Subsequent manipulation with zygotes and raising the offspring is not complicated and the similarity of gar and sturgeon eggs should make injecting relatively straightforward. One potentially confounding factor worth mentioning is that gar embryos, like those of lamprey, lack maternal pigmentation, making it difficult to discern when cleavage begins (Comabella et al., 2014).

While other non-teleost ray-finned fish species are likely well suited to CRISPR/Cas9 F0 mutagenesis, two key groups of early-diverging vertebrates, the chondrichthyans, and hagfish, are unlikely to be amenable to this method, at least as described here. Representatives of both groups have been used in developmental studies in the last 2 decades and have provided some fundamental insights into the developmental bases of vertebrate evolution (Ota et al., 2008; Gillis et al., 2009; Oisi et al., 2013; Gillis and Tidswell, 2017). Embryos of both chondrichthyan clades; elasmobranchs (sharks, rays, and skates) and holocephalans (chimaeras) can be collected seasonally in the wild or in public aquaria, sometimes in the dozens (Gillis, 2011). However, zygotes that could potentially be injected are extremely difficult to obtain and, so far, have not been collected in the numbers needed for reproducible gene disruption. The logistics of hagfish embryology are even more challenging. The few classical reports of hagfish embryology have been based on the scarce embryonic material collected in late 19th century. Only recently have substantial numbers of live hagfish embryos been collected and analyzed (Ota and Kuratani, 2006; Kuratani and Ota, 2008; Ota and Kuratani, 2008). To our knowledge, none of these embryos have ever been successfully injected as zygotes.

CONCLUSION

Based on the results reported here, and previous experiences of ourselves and others, we posit that our CRISPR/Cas9 mutagenesis strategy will work well in any vertebrate with zygotes that are easily accessed and can tolerate microinjection. In such animals, we expect the main variable determining success will be injection droplet size, which will likely need to be adjusted according to the egg size and yolk distribution. Current limitations seen in chondrichthyans may be circumvented with techniques that allow mass transfection of cells at later embryonic stages, such as electroporation or viral vectors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by The Animal Research Committee of Research Institute of Fish Culture and Hydrobiology, Faculty of Fisheries and Protection of Waters, University of South Bohemia in České Budějovice, Vodňany, Czech Republic and Ministry of Agriculture of the Czech Republic (MSMT-12550/2016-3).

AUTHOR CONTRIBUTIONS

DJ conceived and designed the project, JS, VS, RF, AP, VP, MP, and DJ performed the experiments, JS and DJ collected and interpreted the data, DJ, DM, and JS wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

FUNDING

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 897949 (to JS). VS was supported by Charles University Research Centre program No. 204069, Charles University Grant SVV 260571/202, and the grant of the Czech Science Foundation GACR 18-04580S. The work of RF and MP was supported by the Ministry of Education, Youth and Sports of the Czech Republic, projects CENAKVA (LM2018099), Biodiversity (CZ.02.1.01/0.0/0.0/16_025/0007370) and Czech Science Foundation (20-23836S). RC was supported by the Czech Science Foundation GACR 19-18634S. MB was supported by National Institutes of Health grant R35NS111564. DM was supported by National Institutes of Health grant NIDCR R21 RDE025940A and National Science Foundation grant IOS 1656843. DJ was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 751066 and by the Scientific Grant Agency of the Slovak Republic VEGA grant no. 1/0450/21.

ACKNOWLEDGMENTS

We thank Marek Rodina, David Gela, Michaela Fučíková, and Martin Kahanec for their essential help with sterlet spawns; Vojtěch Miller and Štěpánka Novotná for technical assistance; Radek Šanda for allowing us to use the stereoscopic microscope with Z-stacking. Computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures. D.J. thanks Tyler Square for numerous discussions and priceless advice on everything CRISPR-related and Brent Hawkins for late night fish conversations and insightful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.750833/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

Askary, A., Smeeton, J., Paul, S., Schindler, S., Braasch, I., Ellis, N. A., et al. (2016). Ancient Origin of Lubricated Joints in Bony Vertebrates. *Elife* 5, e16415. doi:10.7554/eLife.16415

Baloch, A. R., Franěk, R., Tichopád, T., Fučíková, M., Rodina, M., and Pšenička, M. (2019). Dnd1 Knockout in Sturgeons by CRISPR/Cas9 Generates Germ Cell Free Host for Surrogate Production. *Animals* 9, 174. doi:10.3390/ani9040174

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR Provides Acquired Resistance against Viruses in Prokaryotes. *Science* 315, 1709–1712. doi:10.1126/science.1138140

Barske, L., Fabian, P., Hirschberger, C., Jandzik, D., Square, T., Xu, P., et al. (2020). Evolution of Vertebrate Gill Covers via Shifts in an Ancient Pou3f3 Enhancer. *Proc. Natl. Acad. Sci. USA* 117, 24876–24884. doi:10.1073/pnas.2011531117

Bassett, A. R., Tibbit, C., Ponting, C. P., and Liu, J.-L. (2013). Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cel Rep.* 4, 220–228. doi:10.1016/j.celrep.2013.06.020

Bemis, W. E., Findeis, E. K., and Grande, L. (1997). An Overview of Acipenseriformes. *Environ. Biol. Fishes* 48, 25–71. doi:10.1023/A:1007370213924

Bi, X., Wang, K., Yang, L., Pan, H., Jiang, H., Wei, Q., et al. (2021). Tracing the Genetic Footprints of Vertebrate landing in Non-teleost ray-finned Fishes. *Cell* 184, 1377–1391. doi:10.1016/j.cell.2021.01.046

Blitz, I. L., Biesinger, J., Xie, X., and Cho, K. W. Y. (2013). Biallelic Genome Modification in F0Xenopus Tropicalisembryos Using the CRISPR/Cas System. *Genesis* 51, 827–834. doi:10.1002/dvg.22719

Braasch, I., Gehrke, A. R., Smith, J. J., Kawasaki, K., Manousaki, T., Pasquier, J., et al. (2016). The Spotted Gar Genome Illuminates Vertebrate Evolution and Facilitates Human-Teleost Comparisons. *Nat. Genet.* 48 (4), 427–437. doi:10.1038/ng.3526

Braasch, I., Guiguen, Y., Loker, R., Letaw, J. H., Ferrara, A., Bobe, J., et al. (2014). Connectivity of Vertebrate Genomes: Paired-Related Homeobox (Prrx) Genes in Spotted Gar, Basal Teleosts, and Tetrapods. *Comp. Biochem. Physiol. C: Toxicol. PharmacologyCBP* 163, 24–36. doi:10.1016/j.cbpc.2014.01.005

Brand, M., Granato, M., and Nösslein-Volhard, C. (2002). "Keeping and Raising Zebrafish," in *Zebrafish - A Practical Approach*. Editors C. Nüsslein-Volhard and R. Dahm (Oxford, UK: Oxford University Press).

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., et al. (2016). Maximizing Mutagenesis with Solubilized CRISPR-Cas9 Ribonucleoprotein Complexes. *Development* 143, 2025–2037. doi:10.1242/dev.134809

Chebanov, M. S., and Galich, E. V. (2011). Sturgeon hatchery manual. FAO Fisheries and Aquaculture Technical Paper No. 558. Ankara: FAO, 303.

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., et al. (2013). Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* 155, 1479–1491. doi:10.1016/j.cell.2013.12.001

Chen, J., Wang, W., Tian, Z., Dong, Y., Dong, T., Zhu, H., et al. (2018). Efficient Gene Transfer and Gene Editing in Sterlet (*Acipenser ruthenus*). *Front. Genet.* 9, 117. doi:10.3389/fgene.2018.00117

Comabella, Y., Canabal, J., Hurtado, A., and García-Galano, T. (2014). Embryonic Development of Cuban Gar (Atractosteus Tristoechus) under Laboratory Conditions. *Anat. Histol. Embryol.* 43, 495–502. doi:10.1111/ahe.12101

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339, 819–823. doi:10.1126/science.1231143

Cooper, M. S., and Virta, V. C. (2007). Evolution of Gastrulation in the ray-finned (Actinopterygian) Fishes. *J. Exp. Zool.* 308B, 591–608. doi:10.1002/jez.b.21142

Crawford, K., Diaz Quiroz, J. F., Koenig, K. M., Ahuja, N., Albertin, C. B., and Rosenthal, J. J. C. (2020). Highly Efficient Knockout of a Squid Pigmentation Gene. *Curr. Biol.* 30, 3484–3490. e4. doi:10.1016/j.cub.2020.06.099

Davis, M. C., Dahn, R. D., and Shubin, N. H. (2007). An Autopodial-like Pattern of Hox Expression in the Fins of a Basal Actinopterygian Fish. *Nature* 447, 473–476. doi:10.1038/nature05838

Dettlaff, T. A., Ginsburg, A. S., and Schmalhausen, O. I. (1993). *Sturgeon Fishes—Developmental Biology and Aquaculture*. New York: Springer.

Deveau, H., Garneau, J. E., and Moineau, S. (2010). CRISPR/Cas System and its Role in Phage-Bacteria Interactions. *Annu. Rev. Microbiol.* 64, 475–493. doi:10.1146/annurev.micro.112408.134123

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., et al. (2008). Heritable Targeted Gene Disruption in Zebrafish Using Designed Zinc-finger Nucleases. *Nat. Biotechnol.* 26, 702–708. doi:10.1038/nbt1409

Du, K., Stöck, M., Kneitz, S., Klopp, C., Woltering, J. M., Adolfi, M. C., et al. (2020). The Sterlet sturgeon Genome Sequence and the Mechanisms of Segmental Rediploidization. *Nat. Ecol. Evol.* 4, 841–852. doi:10.1038/s41559-020-1166-x

Faruqi, A. F., Egholm, M., and Glazer, P. M. (1998). Peptide Nucleic Acid-Targeted Mutagenesis of a Chromosomal Gene in Mouse Cells. *Proc. Natl. Acad. Sci.* 95, 1398–1403. doi:10.1073/pnas.95.4.1398

Fei, J.-F., Schuez, M., Tazaki, A., Taniguchi, Y., Roensch, K., and Tanaka, E. M. (2014). CRISPR-mediated Genomic Deletion of Sox2 in the Axolotl Shows a Requirement in Spinal Cord Neural Stem Cell Amplification during Tail Regeneration. *Stem Cel Rep.* 3, 444–459. doi:10.1016/j.stemcr.2014.06.018

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency Off-Target Mutagenesis Induced by CRISPR-Cas Nucleases in Human Cells. *Nat. Biotechnol.* 31, 822–826. doi:10.1038/nbt.2623

Funk, E. C., Breen, C., Sanketi, B. D., Kurpios, N., and McCune, A. (2020). Changes in Nkx2.1, Sox2, Bmp4 , and Bmp16 Expression Underlying the Lung-to-gas Bladder Evolutionary Transition in ray-finned Fishes. *Evol. Develop.* 22, 384–402. doi:10.1111/ede.12354

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas Bacterial Immune System Cleaves Bacteriophage and Plasmid DNA. *Nature* 468, 67–71. doi:10.1038/nature09523

Gillis, J. A., Dahn, R. D., and Shubin, N. H. (2009). Shared Developmental Mechanisms Pattern the Vertebrate Gill Arch and Paired Fin Skeletons. *Proc. Natl. Acad. Sci.* 106, 5720–5724. doi:10.1073/pnas.0810959106

Gillis, J. A. (2011). Hard-to-find Fish Reveals Shared Developmental Toolbox of Evolution Elephant Fish. Available at: https://www.cam.ac.uk/research/news/hard-to-find-fish-reveals-shared-developmental-toolbox-of-evolution (Accessed July 15, 2021).

Gillis, J. A., and Tidswell, O. R. A. (2017). The Origin of Vertebrate Gills. *Curr. Biol.* 27, 729–732. doi:10.1016/j.cub.2017.01.022

Havelka, M., Bytyutskyy, D., Symonová, R., Ráb, P., and Flajšhans, M. (2016). The Second Highest Chromosome Count Among Vertebrates Is Observed in Cultured sturgeon and Is Associated with Genome Plasticity. *Genet. Sel Evol.* 48, 12. doi:10.1186/s12711-016-0194-0

Havelka, M., Hulák, M., Bailie, D. A., Prodöhl, P. A., and Flajšhans, M. (2013). Extensive Genome Duplications in Sturgeons: New Evidence from Microsatellite Data. *J. Appl. Ichthyol.* 29, 704–708. doi:10.1111/jai.12224

Hawkins, M. B., Henke, K., and Harris, M. P. (2021). Latent Developmental Potential to Form Limb-like Skeletal Structures in Zebrafish. *Cell* 184, 899–911. e13. doi:10.1016/j.cell.2021.01.003

Hedges, S. B. (2002). The Origin and Evolution of Model Organisms. *Nat. Rev. Genet.* 3, 838–849. doi:10.1038/nrg929

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science* 327, 167–170. doi:10.1126/science.1179555

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157, 1262–1278. doi:10.1016/j.cell.2014.05.010

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA Targeting Specificity of RNA-Guided Cas9 Nucleases. *Nat. Biotechnol.* 31, 827–832. doi:10.1038/nbt.2647

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., and Zhang, B. (2011). Heritable Gene Targeting in Zebrafish Using Customized TALENs. *Nat. Biotechnol.* 29, 699–700. doi:10.1038/nbt.1939

Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., et al. (2018). Comprehensive Phylogeny of ray-finned Fishes (Actinopterygii) Based on Transcriptomic and Genomic Data. *Proc. Natl. Acad. Sci. USA* 115, 6249–6254. doi:10.1073/pnas.1719358115

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System. *Nat. Biotechnol.* 31 (3), 227–229. doi:10.1038/nbt.2501

Jansen, G., Hazendonk, E., Thijssen, K. L., and Plasterk, R. H. A. (1997). Reverse Genetics by Chemical Mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* 17, 119–121. doi:10.1038/ng0997-119

Jao, L.-E., Wente, S. R., and Chen, W. (2013). Efficient Multiplex Biallelic Zebrafish Genome Editing Using a CRISPR Nuclease System. *Proc. Natl. Acad. Sci. USA* 110, 13904–13909. doi:10.1073/pnas.1308335110

Jao, L.-E., Wente, S. R., and Chen, W. (2013). Efficient Multiplex Biallelic Zebrafish Genome Editing Using a CRISPR Nuclease System. *Proc. Natl. Acad. Sci. USA* 110, 13904–13909. doi:10.1073/pnas.1308335110

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Kaiser, K., and Goodwin, S. F. (1990). "Site-selected" Transposon Mutagenesis of Drosophila. *Proc. Natl. Acad. Sci.* 87, 1686–1690. doi:10.1073/pnas.87.5.1686

Kiyonari, H., Kaneko, M., Abe, T., Shiraishi, A., Yoshimi, R., Inoue, K.-i., et al. (2021). Targeted Gene Disruption in a Marsupial, Monodelphis Domestica, by CRISPR/Cas9 Genome Editing. *Curr. Biol.* 31, 3956–3963. doi:10.1016/j.cub.2021.06.056

Komor, A. C., Badran, A. H., and Liu, D. R. (2017). CRISPR-based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* 168, 20–36. doi:10.1016/j.cell.2016.10.044

Kuratani, S., and Ota, K. G. (2008). Hagfish (Cyclostomata, Vertebrata): Searching for the Ancestral Developmental Plan of Vertebrates. *Bioessays* 30, 167–172. doi:10.1002/bies.20701

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and Classification of the CRISPR-Cas Systems. *Nat. Rev. Microbiol.* 9, 467–477. doi:10.1038/nrmicro2577

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., et al. (2013). RNA-guided Human Genome Engineering via Cas9. *Science* 339, 823–826. doi:10.1126/science.1232033

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., et al. (2016). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* 26, 14–26. doi:10.1016/j.cub.2015.11.021

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D., and Wolfe, S. A. (2008). Targeted Gene Inactivation in Zebrafish Using Engineered Zinc-finger Nucleases. *Nat. Biotechnol.* 26, 695–701. doi:10.1038/nbt1398

Minarik, M., Stundl, J., Fabian, P., Jandzik, D., Metscher, B. D., Psenicka, M., et al. (2017). Pre-oral Gut Contributes to Facial Structures in Non-teleost Fishes. *Nature* 547 (7662), 209–212. doi:10.1038/nature23008

Mizuno, S., Dinh, T. T. H., Kato, K., Mizuno-Iijima, S., Tanimoto, Y., Daitoku, Y., et al. (2014). Simple Generation of Albino C57BL/6J Mice with G291T Mutation in the Tyrosinase Gene by the CRISPR/Cas9 System. *Mamm. Genome* 25, 327–334. doi:10.1007/s00335-014-9524-0

Modrell, M. S., Bemis, W. E., Northcutt, R. G., Davis, M. C., and Baker, C. V. H. (2011). Electrosensory Ampullary Organs Are Derived from Lateral Line Placodes in Bony Fishes. *Nat. Commun.* 2, 496. doi:10.1038/ncomms1502

Modrell, M. S., Lyne, M., Carr, A. R., Zakon, H. H., Buckley, D., Campbell, A. S., et al. (2017). Insights into Electrosensory Organ Development, Physiology and Evolution from a Lateral Line-Enriched Transcriptome. *eLife* 6, e24197. doi:10.7554/eLife.24197

Mori, S., and Nakamura, T. (2021). An Evolutionarily Conserved Odontode Gene Regulatory Network Underlies Head Armor Formation in Suckermouth Armored Catfish. *bioRxiv*. 54. doi:10.1101/2021.06.21.449322

Nakayama, T., Fish, M. B., Fisher, M., Oomen-Hajagos, J., Thomsen, G. H., and Grainger, R. M. (2013). Simple and Efficient CRISPR/Cas9-mediated Targeted Mutagenesis inXenopus Tropicalis. *Genesis* 51, 835–843. doi:10.1002/dvg.22720

Oisi, Y., Ota, K. G., Kuraku, S., Fujimoto, S., and Kuratani, S. (2013). Craniofacial Development of Hagfishes and the Evolution of Vertebrates. *Nature* 493, 175–180. doi:10.1038/nature11794

Ostaszewska, T., and Dabrowski, K. (2009). "Early Development of Acipenseriformes (Chondrostei", in *Development of Non-teleost Fishes*, ed by Y. W. Kunz, C. A. Luer, and B. G. Kapoor (New Hampshire: CRC Press), 170–229.

Ota, K. G., Kuraku, S., and Kuratani, S. (2008). Hagfish Embryology with Reference to the Evolution of the Neural Crest. *Nature* 446, 672–675. doi:10.1038/nature05633

Ota, K. G., and Kuratani, S. (2008). Developmental Biology of Hagfishes, with a Report on Newly Obtained Embryos of the Japanese Inshore Hagfish, *Eptatretus burgeri*. *Zoolog. Sci.* 25, 999–1011. doi:10.2108/zsj.25.999

Ota, K. G., and Kuratani, S. (2006). The History of Scientific Endeavors towards Understanding Hagfish Embryology. *Zoolog. Sci.* 23, 403–418. doi:10.2108/zsj.23.403

Port, F., Chen, H.-M., Lee, T., and Bullock, S. L. (2014). Optimized CRISPR/Cas Tools for Efficient Germline and Somatic Genome Engineering in Drosophila. *Proc. Natl. Acad. Sci.* 111, E2967–E2976. doi:10.1073/pnas.1405500111

Rajkov, J., Shao, Z., and Berrebi, P. (2014). Evolution of Polyploidy and Functional Diploidization in Sturgeons: Microsatellite Analysis in 10 Sturgeon Species. *J. Hered.* 105, 521–531. doi:10.1093/jhered/esu027

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome Engineering Using the CRISPR-Cas9 System. *Nat. Protoc.* 8, 2281–2308. doi:10.1038/nprot.2013.143

Rasys, A. M., Park, S., Ball, R. E., Alcala, A. J., Lauderdale, J. D., and Menke, D. B. (2019). CRISPR-Cas9 Gene Editing in Lizards through Microinjection of Unfertilized Oocytes. *Cel Rep.* 28, 2288–2292. e3. doi:10.1016/j.celrep.2019.07.089

Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K., and Engels, W. R. (1988). A Stable Genomic Source of P Element Transposase in *Drosophila melanogaster*. *Genetics* 118, 461–470. doi:10.1093/genetics/118.3.461

Saito, T., Pšenička, M., Goto, R., Adachi, S., Inoue, K., Arai, K., et al. (2014). The Origin and Migration of Primordial Germ Cells in Sturgeons. *PLoS One* 9, e86861. doi:10.1371/journal.pone.0086861

Sive, H. L., Grainger, R. M., and Harland, R. M. (2000). *Early Development of* Xenopus laevis*: A Laboratory Manual*. NY: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Square, T. A., Jandzik, D., Massey, J. L., Romášek, M., Stein, H. P., Hansen, A. W., et al. (2020). Evolution of the Endothelin Pathway Drove Neural Crest Cell Diversification. *Nature* 585, 563–568. doi:10.1038/s41586-020-2720-z

Square, T., Romášek, M., Jandzik, D., Cattell, M. V., Klymkowsky, M., and Medeiros, D. M. (2015). CRISPR/Cas9-mediated Mutagenesis in the Sea Lamprey, *Petromyzon marinus*: a Powerful Tool for Understanding Ancestral Gene Functions in Vertebrates. *Development (Cambridge, England)* 142 (23), 4180–4187. doi:10.1242/dev.125609

Stolfi, A., Gandhi, S., Salek, F., and Christiaen, L. (2014). Tissue-specific Genome Editing in Ciona Embryos by CRISPR/Cas9. *Development* 141, 4115–4120. doi:10.1242/dev.114488

Stundl, J., Pospisilova, A., Jandzik, D., Fabian, P., Dobiasova, B., Minarik, M., et al. (2019). Bichir External Gills Arise via Heterochronic Shift that Accelerates Hyoid Arch Development. *Elife* 8, 1–13. doi:10.7554/eLife.43531

Stundl, J., Pospisilova, A., Matějková, T., Psenicka, M., Bronner, M. E., and Cerny, R. (2020). Migratory Patterns and Evolutionary Plasticity of Cranial Neural Crest Cells in ray-finned Fishes. *Develop. Biol.* 467 (1-2), 14–29. doi:10.1016/j.ydbio.2020.08.007

Suzuki, M., Hayashi, T., Inoue, T., Agata, K., Hirayama, M., Suzuki, M., et al. (2018). Cas9 Ribonucleoprotein Complex Allows Direct and Rapid Analysis of Coding and Noncoding Regions of Target Genes in *Pleurodeles waltl* Development and Regeneration. *Develop. Biol.* 443, 127–136. doi:10.1016/j.ydbio.2018.09.008

Symonová, R., Havelka, M., Amemiya, C. T., Howell, W. M., Kořínková, T., Flajšhans, M., et al. (2017). Molecular Cytogenetic Differentiation of Paralogs of Hox Paralogs in Duplicated and Re-diploidized Genome of the North American Paddlefish (*Polyodon spathula*). *BMC Genet.* 18, 1–12. doi:10.1186/s12863-017-0484-8

Takeuchi, M., Takahashi, M., Okabe, M., and Aizawa, S. (2009). Germ Layer Patterning in Bichir and Lamprey; an Insight into its Evolution in Vertebrates. *Develop. Biol.* 332, 90–102. doi:10.1016/j.ydbio.2009.05.543

Thompson, A. W., Hawkins, M. B., Parey, E., Wcisel, D. J., Ota, T., Kawasaki, K., et al. (2021). The Bowfin Genome Illuminates the Developmental Evolution of Ray-Finned Fishes. *Nat. Genet.* 53 (9), 1373–1384. doi:10.1038/s41588-021-00914-y

Trible, W., Olivos-Cisneros, L., McKenzie, S. K., Saragosti, J., Chang, N.-C., Matthews, B. J., et al. (2017). Orco Mutagenesis Causes Loss of Antennal Lobe Glomeruli and Impaired Social Behavior in Ants. *Cell* 170, 727–735. e10. doi:10.1016/j.cell.2017.07.001

Véron, N., Qu, Z., Kipen, P. A. S., Hirst, C. E., and Marcelle, C. (2015). CRISPR Mediated Somatic Cell Genome Engineering in the Chicken. *Develop. Biol.* 407, 68–74. doi:10.1016/j.ydbio.2015.08.007

Waaijers, S., Portegijs, V., Kerver, J., Lemmens, B. B. L. G., Tijsterman, M., van den Heuvel, S., et al. (2013). CRISPR/Cas9-targeted Mutagenesis in *Caenorhabditis elegans*. *Genetics* 195, 1187–1191. doi:10.1534/genetics.113.156299

Wang, F., Shi, Z., Cui, Y., Guo, X., Shi, Y.-B., and Chen, Y. (2015). Targeted Gene Disruption in *Xenopus laevis* Using CRISPR/Cas9. *Cell Biosci* 5, 15. doi:10.1186/s13578-015-0006-1

Wood, A. J., Lo, T.-W., Zeitler, B., Pickle, C. S., Ralston, E. J., Lee, A. H., et al. (2011). Targeted Genome Editing across Species Using ZFNs and TALENs. *Science* 333, 307. doi:10.1126/science.1207773

Yen, S.-T., Zhang, M., Deng, J. M., Usman, S. J., Smith, C. N., Parker-thornburg, J., et al. (2014). Somatic Mosaicism and Allele Complexity Induced by CRISPR/Cas9 RNA Injections in Mouse Zygotes. *Develop. Biol.* 393, 3–9. doi:10.1016/j.ydbio.2014.06.017

York, J. R., Yuan, T., Zehnder, K., and McCauley, D. W. (2017). Lamprey Neural Crest Migration Is Snail-dependent and Occurs without a Differential Shift in Cadherin Expression. *Develop. Biol.* 428 (1), 176–187. doi:10.1016/j.ydbio.2017.06.002

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Stundl, Soukup, Franěk, Pospisilova, Psutkova, Pšenička, Cerny, Bronner, Medeiros and Jandzik. 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) and the copyright owner(s) 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
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership