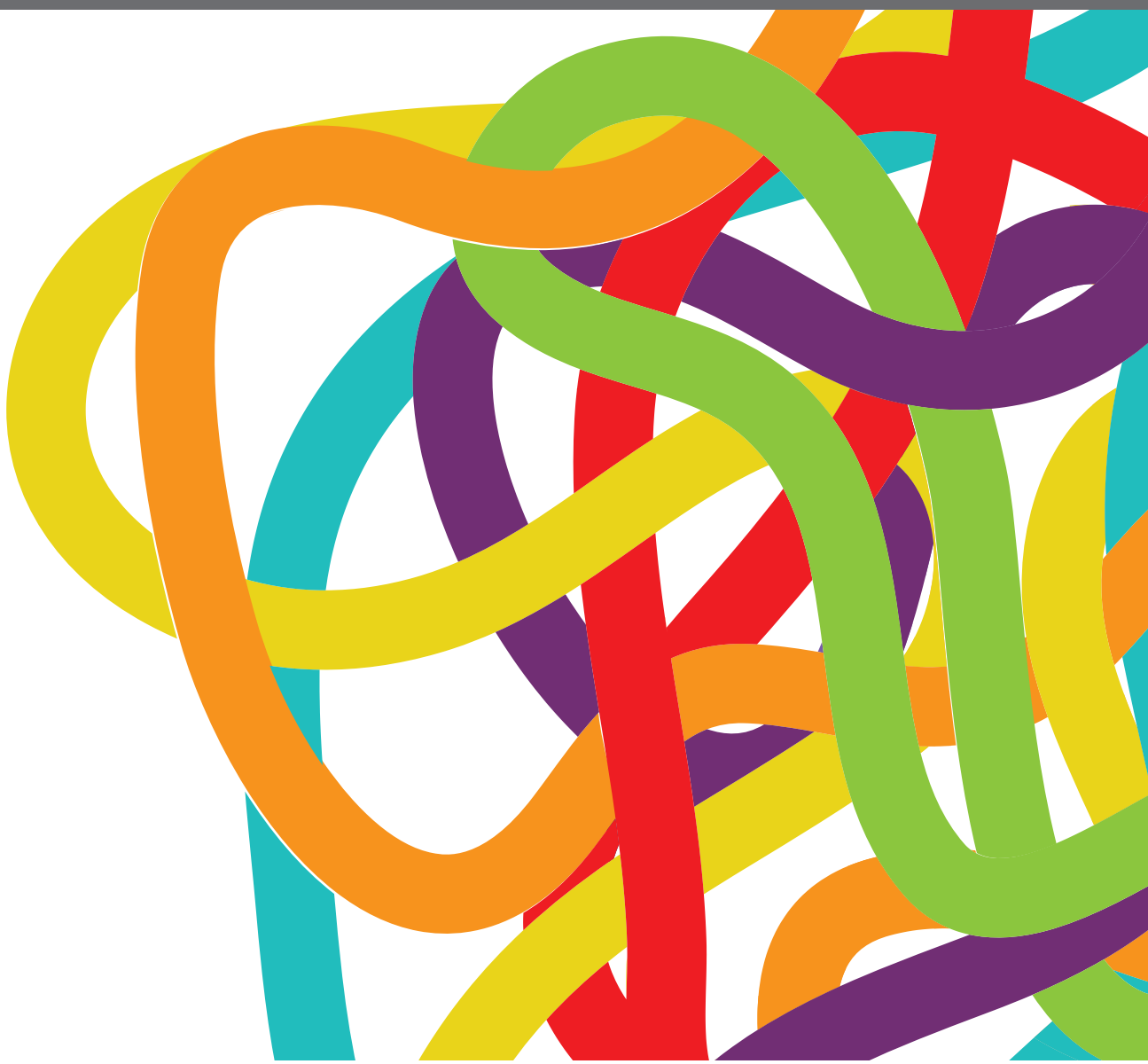
PROTEOMICS AND ITS APPLICATIONS IN CANCER

EDITED BY: Suman S. Thakur, Harsha Gowda, D. R. Mani and Bhaswati Chatterjee
PUBLISHED IN: Frontiers in Oncology and
Frontiers in Cell and Developmental Biology



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 978-2-88971-762-0
DOI 10.3389/978-2-88971-762-0

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

PROTEOMICS AND ITS APPLICATIONS IN CANCER

Topic Editors:
Suman S. Thakur, Centre for Cellular & Molecular Biology (CCMB), India
Harsha Gowda, The University of Queensland, Australia
D. R. Mani, Broad Institute, United States
Bhaswati Chatterjee, National Institute of Pharmaceutical Education and Research, India

Citation: Thakur, S. S., Gowda, H., Mani, D. R., Chatterjee, B., eds. (2021). Proteomics and its Applications in Cancer. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-762-0

Table of Contents

06 ***Editorial: Proteomics and Its Applications in Cancer***
Suman S. Thakur, Harsha Gowda, D. R. Mani and Bhaswati Chatterjee

09 ***WDHD1 Leads to Cisplatin Resistance by Promoting MAPRE2 Ubiquitination in Lung Adenocarcinoma***
Lian Gong, Mengqing Xiao, Dong He, Yi Hu, Yuxing Zhu, Liang Xiang, Ying Bao, Xiaoming Liu, Qinghai Zeng, Jianye Liu, Ming Zhou, Yanhong Zhou, Yaxin Cheng, Yeyu Zhang, Liping Deng, Rongrong Zhu, Hua Lan and Ke Cao

23 ***Proteomic and Phosphoproteomic Maps of Lung Squamous Cell Carcinoma From Chinese Patients***
Lulu Pan, Xijun Wang, Longhai Yang, Lei Zhao, Linhui Zhai, Junyu Xu, Yikun Yang, Yousheng Mao, Shujun Cheng, Ting Xiao and Minjia Tan

36 ***Proteomic Analysis of the Air-Way Fluid in Lung Cancer. Detection of Periostin in Bronchoalveolar Lavage (BAL)***
Yangying Zhou, Weiming Yang, Minghui Ao, Naseruddin Höti, Edward Gabrielson, Daniel W. Chan, Hui Zhang and Qing Kay Li

45 ***Biliary Tract Carcinogenesis Model Based on Bile Metaproteomics***
Ariel A. Arteta, Miryan Sánchez-Jiménez, Diego F. Dávila, Oscar G. Palacios and Nora Cardona-Castro

60 ***Schwann Cell Stimulation of Pancreatic Cancer Cells: A Proteomic Analysis***
Aysha Ferdoushi, Xiang Li, Nathan Griffin, Sam Faulkner, M. Fairuz B. Jamaluddin, Fangfang Gao, Chen Chen Jiang, Dirk F. van Helden, Pradeep S. Tanwar, Phillip Jobling and Hubert Hondermarck

74 ***Mutant Proteomics of Lung Adenocarcinomas Harboring Different* EGFR *Mutations***
Toshihide Nishimura, Ákos Végvári, Haruhiko Nakamura, Harubumi Kato and Hisashi Saji

94 ***Comprehending Meningioma Signaling Cascades Using Multipronged Proteomics Approaches & Targeted Validation of Potential Markers***
Shuvolina Mukherjee, Deeptarup Biswas, Rucha Gadre, Pooja Jain, Nelofer Syed, Julianna Stylianou, Qingyu Zeng, Anita Mahadevan, Sridhar Epari, Prakash Shetty, Aliasgar Moiyadi, Graham Roy Ball and Sanjeeva Srivastava

110 ***Toxicoproteomics Disclose Pesticides as Downregulators of TNF-α, IL-1β and Estrogen Receptor Pathways in Breast Cancer Women Chronically Exposed***
Luciana Pizzatti, Aedra Carla Bufalo Kawassaki, Bruna Fadel, Fabio C. S. Nogueira, Joseph A. M. Evaristo, Nicole Woldmar, Géssica Tuani Teixeira, Janaína Carla Da Silva, Thalita Basso Scandolara, Daniel Rech, Luciano Pessôa Zanetti Candiotto, Guilherme Ferreira Silveira, Wander Rogério Pavanelli and Carolina Panis

121 ***FBXO2 Promotes Proliferation of Endometrial Cancer by Ubiquitin-Mediated Degradation of FBN1 in the Regulation of the Cell Cycle and the Autophagy Pathway***
Xiaoxia Che, Fangfang Jian, Ying Wang, Jingjing Zhang, Jian Shen, Qi Cheng, Xi Wang, Nan Jia and Weiwei Feng

136 ***NEDD4 Plays Roles in the Maintenance of Breast Cancer Stem Cell Characteristics***
Seon-Ae Jeon, Dong Wook Kim, Da-Bin Lee and Je-Yoel Cho

148 ***An Integrated Quantitative Proteomics Workflow for Cancer Biomarker Discovery and Validation in Plasma***
Vipin Kumar, Sandipan Ray, Saicharan Ghantasala and Sanjeeva Srivastava

162 ***Ovarian Blood Sampling Identifies Junction Plakoglobin as a Novel Biomarker of Early Ovarian Cancer***
Florian Weiland, Noor A. Lokman, Manuela Klingler-Hoffmann, Thomas Jobling, Andrew N. Stephens, Karin Sundfeldt, Peter Hoffmann and Martin K. Oehler

172 ***Down-Regulation of Collagen Hydroxylation in Colorectal Liver Metastasis***
Nick A. van Huizen, Peter C. Burgers, Joost van Rosmalen, Michail Doukas, Jan N. M. IJzermans and Theo M. Luider

183 ***Proteomic Profiling of Plasma and Total Blood Exosomes in Breast Cancer: A Potential Role in Tumor Progression, Diagnosis, and Prognosis***
Oleg Tutanov, Ksenia Proskura, Roman Kamyshinsky, Tatiana Shtam, Yuri Tsentalovich and Svetlana Tamkovich

199 ***Comparative Assessment of the WNT/β-Catenin Pathway, CacyBP/SIP, and the Immunoproteasome Subunit LMP7 in Various Histological Types of Renal Cell Carcinoma***
Żaneta Piotrowska, Michał Niezgoda, Grzegorz Młynarczyk, Magdalena Acewicz and Irena Kasacka

211 ***Dissection of Anti-tumor Activity of Histone Deacetylase Inhibitor SAHA in Nasopharyngeal Carcinoma Cells* via *Quantitative Phosphoproteomics***
Huichao Huang, Ying Fu, Ye Zhang, Fang Peng, Miaolong Lu, Yilu Feng, Lin Chen, Zhuchu Chen, Maoyu Li and Yongheng Chen

227 ***Proteome-Wide Alterations of Asymmetric Arginine Dimethylation Associated With Pancreatic Ductal Adenocarcinoma Pathogenesis***
Meijin Wei, Chaochao Tan, Zhouqin Tang, Yingying Lian, Ying Huang, Yi Chen, Congwei Chen, Wen Zhou, Tao Cai and Jiliang Hu

241 ***Comparative Analysis of Systemic and Tumor Microenvironment Proteomes From Children With B-Cell Acute Lymphocytic Leukemia at Diagnosis and After Induction Treatment***
Geise Ellen Broto, Stephany Corrêa, Fausto Celso Trigo, Everton Cruz dos Santos, Fernanda Tomiotto-Pelissier, Wander Rogério Pavanelli, Guilherme Ferreira Silveira, Eliana Abdelhay and Carolina Panis

254 ***Proteomic Analysis of Pre-Invasive Serous Lesions of the Endometrium and Fallopian Tube Reveals Their Metastatic Potential***
Mitchell Acland, Georgia Arentz, Max Mussared, Fergus Whitehead, Peter Hoffmann, Manuela Klingler-Hoffmann and Martin K. Oehler

266 *A Protein Microarray-Based Investigation of Cerebrospinal Fluid Reveals Distinct Autoantibody Signature in Low and High-Grade Gliomas*
Nikita Gahoi, Parvez Syed, Saket Choudhary, Sridhar Epari, Aliasgar Moiyadi, Santosh G. Varma, Mayuri N. Gandhi and Sanjeeva Srivastava

274 *KIF21B Expression in Osteosarcoma and Its Regulatory Effect on Osteosarcoma Cell Proliferation and Apoptosis Through the PI3K/AKT Pathway*
Songjia Ni, Jianjun Li, Sujun Qiu, Yingming Xie, Kaiqin Gong and Yang Duan

289 *Proteomic Alterations in Multiple Myeloma: A Comprehensive Study Using Bone Marrow Interstitial Fluid and Serum Samples*
Venkatesh Chanukuppa, Ravindra Taware, Khushman Taunk, Tathagat Chatterjee, Sanjeevan Sharma, Venkatesan Somasundaram, Faraz Rashid, Dipankar Malakar, Manas K. Santra and Srikanth Rapole

306 *Identifying Apoptosis-Related Transcriptomic Aberrations and Revealing Clinical Relevance as Diagnostic and Prognostic Biomarker in Hepatocellular Carcinoma*
Jinyu Zhu, Bufu Tang, Xiuling Lv, Miaomiao Meng, Qiaoyou Weng, Nannan Zhang, Jie Li, Kai Fan, Liyun Zheng, Shiji Fang, Min Xu and Jiansong Ji

321 *Application of Advanced Mass Spectrometry-Based Proteomics to Study Hypoxia Driven Cancer Progression*
Arada Vinaiphat, Jee Keem Low, Kheng Wei Yeoh, Wee Joo Chng and Siu Kwan Sze

334 *Quantitative Proteomic Profiling Identifies a Potential Novel Chaperone Marker in Resistant Breast Cancer*
Karen M. Gallegos, Jankiben R. Patel, Shawn D. Llopis, Rashidra R. Walker, A. Michael Davidson, Wensheng Zhang, Kun Zhang and Syreeta L. Tilghman

349 *Protein and Signaling Pathway Responses to rhIL-6 Intervention Before Lobaplatin Treatment in Osteosarcoma Cells*
Huan Wang, Bin Li, Kang Yan, Yonghong Wu, Yanhua Wen, Yunyan Liu, Pei Fan and Qiong Ma

366 *Transcriptional Repression of MFG-E8 Causes Disturbance in the Homeostasis of Cell Cycle Through DOCK/ZP4/STAT Signaling in Buffalo Mammary Epithelial Cells*
Arvind K. Verma, Syed A. Ali, Parul Singh, Sudarshan Kumar and Ashok K. Mohanty

388 *Multiple Reaction Monitoring-Based Targeted Assays for the Validation of Protein Biomarkers in Brain Tumors*
Saicharan Ghantasala, Medha Gayathri J. Pai, Deeptarup Biswas, Nikita Gahoi, Shuvolina Mukherjee, Manubhai KP, Mehar Un Nissa, Alisha Srivastava, Sridhar Epari, Prakash Shetty, Aliasgar Moiyadi and Sanjeeva Srivastava

399 *Proteomic Characterization of Cytoplasmic Lipid Droplets in Human Metastatic Breast Cancer Cells*
Alyssa S. Zembroski, Chaylen Andolino, Kimberly K. Buhman and Dorothy Teegarden

412 *Coupling of Cell Surface Biotinylation and SILAC-Based Quantitative Proteomics Identified Myoferlin as a Potential Therapeutic Target for Nasopharyngeal Carcinoma Metastasis*
Maoyu Li, Fang Peng, Guoqiang Wang, Xujun Liang, Meiying Shao, Zhuchu Chen and Yongheng Chen

frontiers in Oncology

EDITORIAL
published: 07 October 2021
doi: 10.3389/fonc.2021.772811

Editorial: Proteomics and Its Applications in Cancer

Suman S. Thakur[1], Harsha Gowda[2]*, D. R. Mani[3]* and Bhaswati Chatterjee[4]**

[1] Proteomics and Signaling, Centre for Cellular and Molecular Biology, Hyderabad, India, [2] Queensland Institute of Medical Research (QIMR) Berghofer Medical Research Institute, The University of Queensland, Brisbane, QLD, Australia, [3] Computational Proteomics, The Broad Institute of Massachusetts Institute of Technology (MIT) and Harvard, Cambridge, MA, United States, [4] Chemical Science, National Institute of Pharmaceutical Education and Research, Hyderabad, India

Keywords: proteomics, cancer, post-translational modifications, mass spectrometry, quantitative proteomics, therapeutic, biomarkers

OPEN ACCESS

Edited and reviewed by:
Tao Liu,
University of New South Wales, Australia

****Correspondence:***
Suman S. Thakur
sst@ccmb.res.in
Harsha Gowda
Harsha.Gowda@qimrberghofer.edu.au
D. R. Mani
manidr@broadinstitute.org
Bhaswati Chatterjee
bhaswatichatterjee@yahoo.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *08 September 2021*
Accepted: *17 September 2021*
Published: *07 October 2021*

Citation:
Thakur SS, Gowda H, Mani DR and Chatterjee B (2021) Editorial: Proteomics and Its Applications in Cancer. Front. Oncol. 11:772811. doi: 10.3389/fonc.2021.772811

Editorial on the Research Topic

Proteomics and Its Applications in Cancer

Proteomics has great potential to provide new insights into molecular and mechanistic processes that govern oncogenesis. Unraveling complex signaling mechanisms mediated by various post-translational modifications (PTMs), including phosphorylation, glycosylation, and ubiquitination, is an area where proteomics has the greatest impact. It is a valuable tool that could provide insight into the mode of action of new and established drugs. Mass spectrometry also allows the investigation of drug-protein interactions to enable the repurposing of drugs. High-throughput quantitative proteomics approaches including label-free and label-based quantification methods (SILAC, iTRAQ, and TMT) are extensively used to identify potential biomarkers and therapeutic targets. Here, we highlight some of these applications in studies published in this special issue.

PTMs AND SIGNALING

Post-translational modifications of proteins regulate the biological function of proteins and their role in signaling. Zhou et al detected hundreds of N-glycoproteins in airway fluid bronchoalveolar lavage (BAL) present in lung adenocarcinoma (ADC), squamous cell carcinoma (SQCC), and small cell lung carcinoma (SCLC). Further, a significantly elevated level of periostin was observed in all subtypes of lung cancer samples. Moreover, a lower dimethylation of asymmetric dimethylarginine (ADMA) containing protein was found in pancreatic ductal adenocarcinoma (PDAC) cell line PANC-1 (Wei et al). van Huizen et al reported a significantly lower degree of hydroxylation of collagen alpha-1 peptides in the tissue of colorectal cancer (CRC) and colorectal liver metastasis (CRLM) compared to colon tissues. Moreover, Suberoylanilide hydroxamic acid (SAHA) activates tumor suppressor p53 and Rb1 through phosphorylation and causes apoptosis in nasopharyngeal carcinoma (NPC) cells (Huang et al).

Interestingly, a significant increase of WNT10A, β-catenin, GSK-3ß, LMP7, and CacyBP/SIP were observed in renal cell carcinoma (RCC) compared to the control, suggesting the involvement of WNT/β-catenin pathway in RCC (Piotrowska et al). KIF21B was found to be highly upregulated in osteosarcoma, targeting PI3K/AKT pathway and playing an important role in the regulation of osteosarcoma cell proliferation and apoptosis (Ni et al).

DIAGNOSTIC/PROGNOSTIC BIOMARKERS

Biomarkers can play an important role in the detection of early stage of cancer, recurrence, and treatment to reduce mortality rate. Tutanov et al reported SERPINA1, KRT6B, and SOCS3 as prognostic protein markers in breast cancer patients' total blood exosomes. Proteins related to cholesterol synthesis were identified, including squalene monooxygenase (SQLE) and sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating (NSDHL) in cytoplasmic lipid droplet proteomics of breast cancer cells (Zembroski et al). Further, neural precursor cells expressing developmentally downregulated protein 4 (NEDD4) play an important role in the maintenance of a significantly high number of characteristic cancer stem cells (CSCs) in triple-negative breast cancer (TNBC) (Jeon et al). Another study found the expression of Lactadherin or MFG-E8 in mammary epithelial cells is associated with intracellular deterioration (Verma et al). The junction plakoglobin (JUP) is an important biomarker for epithelial ovarian cancer and, together with Cancer Antigen 125 (CA125), may be used for screening ovarian cancer (Weiland et al). Kumar et al discussed blood cancer biomarker discovery and plasma proteomics using iTRAQ, TMT, MRM, and PRM. Zhu et al reported apoptosis genes BAK1 and CSE1L as potential diagnosis and prognosis markers for hepatocellular carcinoma.

THERAPEUTICS, MODE OF ACTION, AND RESISTANCE

Wang et al studied the effect of lobaplatin in osteosarcoma cells with and without interleukin (IL)-6 using tandem mass tag (TMT) labeling and suggested upregulation of far upstream element-binding protein 1 (FUBP1), fragile X mental retardation syndrome-related protein 1 (hFXR1p), and rasGTPase-activating protein-binding protein 1 (G3BP1) and verified the result by PRM. They further found significantly high expression of FUBP1, hFXR1p, and G3BP1 in platinum-resistant osteosarcoma patients. Interestingly, more active ubiquitin-dependent protein catabolism and higher expression of WDHD1 were observed in the cisplatin-resistant strain A549/DDP as compared to A549 cells. Further, WDHD1 promotes MAPRE2 ubiquitination and causes cisplatin resistance in lung adenocarcinoma (Gong et al).

Gallegos et al carried out proteomics of CSC enriched with letrozole resistant breast cancer cell (LTLT-Ca) mammospheres and reported a significant increase of the chaperone protein midasin which is needed for maturation and nuclear export of the pre-60S ribosome. Nishimura et al attempted to find differences in cells with L858R and Ex19del mutations in epidermal growth factor receptor (EGFR) mutant lung adenocarcinomas to identify the efficacy of tyrosine kinase inhibitors.

INSIGHTS INTO CANCER BIOLOGY

The DNA replication licensing factor MCM2 was significantly expressed in lung squamous cell carcinoma patients (Pan et al). In endometrial carcinoma (EC), the F-box only protein 2 (FBXO2) was significantly upregulated and was related to its tumor stage. Further, the FBXO2 plays an important role as E3 ligase and causes ubiquitin-dependent degradation of Fibrillin1 (FBN1) (Che et al). The upregulation of myoferlin (MYOF), EGFR, and ephrin type-A receptor 2 (EPHA2) were reported in metastatic NPC cells using cell surface biotinylation and SILAC along with the interaction of MYOF with EGFR and EPHA2 (Li et al). Broto et al studied peripheral blood (PB) and bone marrow compartment (BM) samples of B-cell acute lymphocytic leukemia (B-ALL) and suggested the important role of transthyretin and interferon-gamma (IFN-γ) in B-ALL. Notably, the decrease of tumor necrosis factor-alpha (TNF-α) and interleukin 1 beta (IL-1β) was reported in a pesticide-exposed group (Pizzatti et al).

The secretome of Schwann cells (SCs), including matrix metalloproteinase-2, cathepsin D, plasminogen activator inhibitor-1, and galectin-1, may stimulate pancreatic cancer (PC) growth and invasion (Ferdoushi et al). Furthermore, bacterial proteomics of pancreatic head ductal adenocarcinoma (PDAC) suggested dysbiotic microenvironment with the occurrence of surfactin in PDAC and zeatin in gallstone (GS) patients (Arteta et al).

APPLICATION OF PROTEOMICS TECHNOLOGIES

Multiple reaction monitoring was used to quantify peptides from Prostaglandin H2 D-Isomerase (PTGDS), Vitronectin (VTN), and Complement C3 (C3) using CSF from glioma and meningioma patients (Ghantasala et al). The transition of proteins related to meningioma pathobiology, including NEK9 and CKAP4, were observed in targeted proteomics using SRM assay (Mukherjee et al). Further, the changes in bone marrow interstitial fluid (BMIF) and serum proteome in multiple myeloma (MM) patients were reported using various proteomic methods, including SWATH-MS, suggesting the importance of haptoglobin, kininogen 1, and transferrin in MM (Chanukuppa et al). Vinaiphat et al discussed the role of different proteomics approaches in understanding the hypoxia-driven progression of cancer. Acland et al reported a proteomics study of premalignant lesions in endometrial intraepithelial carcinoma (EIC) and serous tubal intraepithelial carcinoma (STIC). Further, Gahoi et al studied cerebrospinal fluid (CSF)

samples of low-grade glioma (LGG) and the glioblastoma multiforme (GBM) patient using protein microarrays.

Thus, this special issue highlights the application of proteomic technologies to gain deeper insight into cancer, with topics spanning biomarker discovery, pathogenesis, mechanistic details of signaling pathways, the effect of post-translational modifications, therapeutic agents and their mode of action, and computational proteomics.

AUTHOR CONTRIBUTIONS

ST, HG, DM, and BC conceived the idea and wrote the manuscript. All authors contributed to the article and approved the submitted version.

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 Thakur, Gowda, Mani and Chatterjee. 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 Oncology

ORIGINAL RESEARCH
published: 24 April 2020
doi: 10.3389/fonc.2020.00461

WDHD1 Leads to Cisplatin Resistance by Promoting MAPRE2 Ubiquitination in Lung Adenocarcinoma

*Lian Gong[1†], Mengqing Xiao[1†], Dong He[2], Yi Hu[1], Yuxing Zhu[1], Liang Xiang[1], Ying Bao[1], Xiaoming Liu[3], Qinghai Zeng[4], Jianye Liu[5], Ming Zhou[6], Yanhong Zhou[6], Yaxin Cheng[1], Yeyu Zhang[1], Liping Deng[1], Rongrong Zhu[1], Hua Lan[7] and Ke Cao[1]**

[1] Department of Oncology, Third Xiangya Hospital of Central South University, Changsha, China, [2] Department of Respiratory, The Second People's Hospital of Hunan Province, Changsha, China, [3] Department of Gastroenterology, Third Xiangya Hospital of Central South University, Changsha, China, [4] Department of Dermatology, Third Xiangya Hospital of Central South University, Changsha, China, [5] Department of Urology, Third Xiangya Hospital of Central South University, Changsha, China, [6] Cancer Research Institute and Key Laboratory of Carcinogenesis of the Chinese Ministry of Health, Central South University, Changsha, China, [7] Department of Gynaecology, Third Xiangya Hospital of Central South University, Changsha, China

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Monica Montopoli, University of Padova, Italy
Hanumantha Rao Madala, Harvard Medical School, United States

***Correspondence:**
Ke Cao
csucaoke@163.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *11 December 2019*
Accepted: *13 March 2020*
Published: *24 April 2020*

Citation:
Gong L, Xiao M, He D, Hu Y, Zhu Y, Xiang L, Bao Y, Liu X, Zeng Q, Liu J, Zhou M, Zhou Y, Cheng Y, Zhang Y, Deng L, Zhu R, Lan H and Cao K (2020) WDHD1 Leads to Cisplatin Resistance by Promoting MAPRE2 Ubiquitination in Lung Adenocarcinoma. Front. Oncol. 10:461. doi: 10.3389/fonc.2020.00461

Ubiquitin ligases have been shown to regulate drug sensitivity. This study aimed to explore the role of the ubiquitin ligase WD repeat and HMG-box DNA binding protein 1 (WDHD1) in regulating cisplatin sensitivity in lung adenocarcinoma (LUAD). A quantitative analysis of the global proteome identified differential protein expression between LUAD A549 cells and the cisplatin-resistant strain A549/DDP. Public databases revealed the relationship between ubiquitin ligase expression and the prognosis of patients with LUAD. Quantitative real-time polymerase chain reaction and Western blotting were used to estimate the WDHD1 expression levels. Analysis of public databases predicted the substrate of WDHD1. Western blotting detected the effect of WDHD1 on microtubule-associated protein RP/EB family member 2 (MAPRE2) and DSTN. Functional analysis of MAPRE2 verified the interaction between WDHD1 and MAPRE2, as well as the interacting sites by methyl-thiazolyl-tetrazolium assay and flow cytometry, immunoprecipitation, protein stability, and immunofluorescence. Cell and animal experiments confirmed the effect of WDHD1 and MAPRE2 on cisplatin sensitivity in LUAD. Clinical data evaluated the impact of WDHD1 expression level on cisplatin sensitivity. Quantitative analysis of the global proteome revealed ubiquitin-dependent protein catabolism to be more active in A549/DDP cells than in A549 cells. WDHD1 expression was higher in A549/DDP cells than in A549 cells, and knocking out WDHD1 increased the sensitivity of A549/DDP cells to cisplatin. WDHD1 overexpression negatively correlated with the overall survival of LUAD patients. We observed that MAPRE2 was upregulated when WDHD1 was knocked out. A MAPRE2 knockout in A549 cells resulted in increased cell viability while decreasing apoptosis when the A549 cells exposed to cisplatin. WDHD1 and MAPRE2 were found to interact in the nucleus, and WDHD1 promoted the ubiquitination of MAPRE2. Following cisplatin exposure, the WDHD1 and MAPRE2 knockout groups facilitated cell proliferation and migration, inhibited apoptosis in A549/DDP cells, decreased apoptosis, and increased tumor size and growth rate in animal experiments.

Immunohistochemistry showed that Ki67 levels increased, and levels of apoptotic indicators significantly decreased in the WDHD1 and MAPRE2 knockout groups. Clinical data confirmed that WDHD1 overexpression negatively correlated with cisplatin sensitivity. Thus, the ubiquitin ligase WDHD1 induces cisplatin resistance in LUAD by promoting MAPRE2 ubiquitination.

Keywords: lung adenocarcinoma, cisplatin, drug sensitivity, ubiquitin ligase, WDHD1, MAPRE2

INTRODUCTION

Lung cancer is the most common type of malignant tumor. It ranks first in morbidity and mortality globally, as well as in China, and is the leading cause of cancer-related deaths (1, 2). The treatment of lung cancer is achieved by a combination of various therapies, including surgery, chemotherapy, radiotherapy, molecular targeted therapy, and immunotherapy. As the early symptoms of lung cancer are not obvious, lung cancer is mostly diagnosed in the advanced stage. Therefore, chemotherapy continues to be an important treatment strategy for lung cancer. The chemotherapy regimen for lung cancer is based on platinum combined with other chemotherapeutic drugs, the most commonly used platinum drug being cisplatin (DDP). However, the failure of lung cancer treatment has typically occurred due to cisplatin resistance. Lung adenocarcinoma (LUAD) accounts for the highest proportion of lung cancer, ranging from 40 to 55%. Therefore, it is critical to explore the molecular mechanisms of cisplatin resistance in LUAD and provide a new basis for cisplatin sensitization in patients with LUAD.

Cisplatin resistance in lung cancer is closely related to DNA damage repair, apoptotic inactivation, activation of epithelial–mesenchymal transition, and characteristics of cancer stem cells (3–8). The related proteins and important molecules in the above signaling pathways are typically regulated by ubiquitination of proteins (9). Ubiquitin is a highly conserved molecule composed of 76 amino acids, which is widely expressed in eukaryotic organisms. The enzymes ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3) together modify a specific target protein by linking ubiquitin to it. E1 activates ubiquitin and transfers it to E2, and E3 recruits ubiquitinated E2, identifies substrates, and assists in the transfer of ubiquitin from E2 to protein substrates (10). E3 plays an important role in recognizing substrates during ubiquitination. Ubiquitin ligases have been shown to be associated with tumor development and a malignant phenotype. Moreover, previous studies have confirmed that E3 mediates cisplatin resistance through the regulation of various substrates.

Overexpression of the ubiquitin ligase MDM2 in malignant pleural mesothelioma is negatively correlated with patient prognosis and leads to p53 degradation and decreased cisplatin sensitivity (11). In addition, the ubiquitin ligase RNF31 is overexpressed in breast cancer tissues and MCF-7 cell lines, and it can promote polyubiquitination and degradation of p53 by stabilizing the ubiquitin ligase MDM2, which results in a reduction in cisplatin-induced apoptosis (12). The ubiquitin ligase NEDD4-1 promotes cisplatin resistance in lung cancer cells by inhibiting PTEN expression and activating Akt and its downstream proteins (13). Hakai (E3) is overexpressed in non–small cell lung cancer cell lines. Hakai interference leads to decreased expression of phosphorylated Akt, which significantly inhibits the growth of non–small cell lung cancer cells and enhances cisplatin drug sensitivity (14). Therefore, ubiquitin ligases can regulate key molecules of the signaling pathway to influence cisplatin resistance.

To explore the causes of cisplatin resistance in patients with LUAD, we screened the protein chip of both A549 and A549/DDP cell lines, as well as identified and analyzed any significantly altered proteins. The protein chip results demonstrated that ubiquitin-dependent protein catabolism was more active in A549/DDP cells than that in A549 cells. On the basis of existing literature and protein chip results, this study aims to investigate the role of ubiquitin ligase in regulating cisplatin sensitivity and its regulatory mechanism in LUAD. Our research may provide a scientific basis and novel insight for reversal of cisplatin resistance and personalized treatment of LUAD in future.

MATERIALS AND METHODS

Cell Culture and Transfection

A549 and A549/DDP cell lines (Central South University, Changsha, China) were cultured in RPMI-1640 cell culture medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal calf serum, 100 U/mL penicillin, and 100 μg/mL streptomycin. The culture medium of the A549/DDP cells contained 2 μg/mL cisplatin (Solarbio Company, Beijing, China), in addition to the other components, and was incubated in 5% CO_2 at 37°C. The purity of cisplatin was 98.5%. All RNA inhibitors and negative control siRNA were purchased from GenePharma (Shanghai, China). The siRNA targeting WD repeat and HMG-box DNA binding protein 1 (WDHD1) (15) (5′-GAUCAGACAUGUGCUAUUA-3′), ARPC1A (16) (5′-GUGGAGCACGACUCAUUUCTT-3′), and microtubule-associated protein RP/EB family member 2 (MAPRE2) (5′-UUGUUC–AGGAGCGGCCUAUTT-3′) were transfected into A549 and A549/DDP cells using Lipofectamine™ 2000 (Invitrogen Life Technologies, Carlsbad, California, USA) according to the manufacturer's instructions. After 24 h, the

Abbreviations: E1, ubiquitin-activating enzyme; E2, ubiquitin-conjugating enzyme; E3, ubiquitin ligase; WDHD1, WD repeat and HMG-box DNA binding protein 1; MAPRE2, microtubule-associated protein RP/EB family member 2; LUAD, lung adenocarcinoma; CO-IP, CO-immunoprecipitation; IHC, immunohistochemical; PCR, polymerase chain reaction; siRNA, small interfering RNA.

supernatant was replaced with culture medium, and the cells were incubated at 37°C in 5% CO_2 for another 24 h.

Quantitative Real-Time Polymerase Chain Reaction

Total RNA was extracted from A549 and A549/DDP cells using TRIzol, and reverse transcription was performed using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Science, Waltham, MA, USA). Real-time polymerase chain reaction (PCR) was performed using SYBR Green Premix PCR Master Mix (Roche, Mannheim, Germany). The relative quantification of WDHD1 and APRC1A was carried out according to the $^{\Delta\Delta}C_T$ method. The primer sequences for amplification of WDHD1 were as follows: forward: 5′-AGGTCGTCCTAGACAGCG-3′; reverse: 5′-GCAT GGGTCCATCATAAA-3′. The primer sequences for ARPC1A were as follows: forward: 5′-CAGTCCCAATAATCACGAA-3′; reverse: 5′-GGAGCCCAGTCAATACCT-3′. β-Actin was used as a housekeeping control. The primer sequences used were as follows: forward: 5′-CATTAAGGAGAAGCTGTGCT-3′; reverse: 5′-GTTGAAGGTAGTTTCGTGG–A-3′ (17).

Western Blotting

The total protein in the cells and tissues was extracted using radioimmunoprecipitation assay (RIPA) buffer (Auragene, Changsha, China). The proteins were quantified using a BCA protein assay kit (Thermo Scientific). The protein samples were separated on a 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gel, followed by transfer of the proteins to polyvinylidene fluoride membranes and blocking of the membranes using 5% Tris buffered saline with Tween-20 buffer containing 5% skim milk. Anti-MAPRE2 (1:100, ab45767; Abcam, Cambridge Science Park, UK), anti-WDHD1 (1:100, ab72436; Abcam), anti-DSTN (1:500, ab186754; Abcam), and anti-ubiquitin antibodies (1:200, 10201-2-AP; Proteintech, Wuhan, China) were used as primary antibodies and incubated with the membrane overnight at 4°C. The membrane was then washed and incubated with a secondary horseradish peroxidase–labeled goat anti–rabbit immunoglobulin G (IgG) antibody (1:4,000; Abcam) for 1 h at room temperature. Immunoassay was performed by enhanced chemiluminescence detection system (ECL; Cell Signaling Technology, Danvers, MA, USA) combined with Western blot system (Auragene). Strip signal strength was checked using the software IPP6.0. GAPDH (1:4,000, ab125247; Abcam) was used as an internal control to normalize the expression of other proteins.

Bioinformatics Analysis

The public database DAVID (https://david.ncifcrf.gov/) was used to analyze the functional enrichment of significantly upregulated proteins. Searching for ubiquitin ligase and proteins interacting with ubiquitin ligase was performed through a website (http://ubibrowser.ncpsb.org/ubibrowser/home/index and http://iuucd.biocuckoo.org/). Public database Gepia, http://gepia.cancer-pku.cn/, was used to find the correlation between gene and prognosis. Searching gene expression was performed on a public database Ualcan (http://ualcan.path.uab.edu/index.html). By referring to the ubiquitination site prediction database Phosphosite, http://www.phosphosite.org, we can find proteins with ubiquitination sites.

Methyl-Thiazolyl-Tetrazolium Assay

Cell viability was assessed using a methyl-thiazolyl-tetrazolium (MTT) assay (18). The A549 and A549/DDP cells were seeded in 96-well plates at a density of 1×10^4 cells per well. The cells were incubated for 24 h following treatment with cisplatin [25% inhibitory concentration (IC25) concentration of cisplatin for A549/DDP cells = 20.240 μg/mL]. Then, 50 μL of 1× MTT (Sigma-Aldrich) was added to each well according to the manufacturer's instructions and incubated at 37°C. After 4 h, MTT solution was removed, and 150 μL dimethyl sulfoxide was added to each well to dissolve the purple formazan crystals. The optical density of each well was measured at a wavelength of 570 nm using a microplate reader (Bio-Rad California, USA). Cell survival rate = OD of the test well/OD of the control well. GraphPad Prism 7.0 was used to create the resulting graph.

Flow Cytometry Assay

Cellular apoptosis was evaluated by flow cytometry. The cells were treated with DDP. To assess apoptosis, an annexin V–[fluorescein isothiocyanate (FITC)]/PI apoptosis detection kit (KeyGEN Biotech, Nanjing, China) was used. After dual staining with FITC–annexin V and propidium iodide the level of fluorescence was measured by flow cytometry (FACS Canto II; BD, New Jersey, USA). Apoptosis rate = early apoptosis rate + late apoptosis rate.

Coimmunoprecipitation

Si-con or si-WDHD1 was transfected into A549/DDP cells. After 24 h, the cells were lysed with RIPA (Auragene). Anti-MAPRE2 (1:100. ab45767; Abcam) and human IgG (1:150, bs-0297P; Bioss, Beijing, China) antibodies were added as precipitation antibodies and incubated overnight at 4°C. Then, 20 μL of protein A + G agarose was added and incubated for 2 h at 4°C. The protein-bead precipitations were denatured in SDS-PAGE by boiling for 5 min at 100°C, and a Western blot was performed to analyze the expression of ubiquitin, WDHD1, and MAPRE2 (19).

Immunofluorescence Staining

The A549/DDP cells were grown until they reached 70% confluency. The cells were fixed with 4% paraformaldehyde stationary solution and incubated for 20 min, washed with phosphate-buffered saline (PBS) three times, and then treated with 1% Triton-x-100 for 10 min. The cells were pretreated with 10% normal goat serum for 30 min. The cells were incubated with the primary antibodies anti-WDHD1 (1:100, ab72436; Abcam) and anti-MAPRE2 (1:100, ab45767; Abcam) antibodies at 4°C overnight and then washed with PBS. The cells were then incubated with an Alexa Fluor® 555 goat anti-rabbit antibody diluted 1:100 in a blocking solution. Nuclei were stained with 4,6-diamidino-2-phenylindole for 15 min, and the cells were examined with Leica TCS-SP5 confocal laser microscope (Heidelberg, Germany).

Wound Healing Assay

Approximately 5×10^5 cells per well were seeded in a 24-well culture plate and treated with cisplatin. After incubation for 24 h,

scratches were made using a sterile 20 μL pipette tip, and the plates were incubated at 37°C in 5% CO_2. After incubation for 0 and 48 h, for each group, three fields with a wound area were chosen randomly and photographed under a microscope. The gap distance was quantitatively evaluated using ImageJ software. Relative mobility = (distance between the edges of migrated scratches/distance between the edges of initial scratches) × 100%.

Colony Formation Assay

The transfected cells were cultured up to the logarithmic growth phase and then trypsinized and seeded into six-well plates (10^3 cells per well). Cisplatin was added to the wells in accordance with the experimental design. Cells were incubated for 2–3 weeks at 37°C, 5% CO_2. Cells were washed with PBS twice before harvest. The cells were fixed in 4% paraformaldehyde for 15 min, stained with hematoxylin, and counted under a microscope. The number of cells contained in each cell clone was counted, and the cell colony formation rate was calculated and photographed. The clone was counted if the cell number of the clone was at least 50. Clonal formation rate = number of clones/number of cells inoculated.

Animal Experiments

All animal experiments were approved by the Animal Laboratory of Central South University and carried out in accordance with international guidelines and programs. Four- to 6-week-old male BALB/C nude mice (purchased from Hunan SJA Laboratory Animal Co., Ltd, Changsha, China) were subcutaneously injected with 1×10^6 cells near the extremities of four limbs. When the mice developed palpable tumors, the mice were intraperitoneally administered with cisplatin (5 mg/kg; Solarbio Company) every week for 2 weeks. The size of the tumors was measured once every 3 days, six times in a row. All animals were sacrificed 25 days after inoculation, and the tumors were collected. The tumor tissues were photographed, and immunohistochemistry was performed. Tumor volume formula: $V = a * b^2 * 0.52$ (mm^3), where a is the longest diameter, and b is the shortest diameter of the tumor.

TUNEL Staining

TUNEL staining was used to detect apoptotic tumor cells (20). The collected tumors were fixed in 4% paraformaldehyde solution for 60 min, embedded in paraffin, and cut into 3-μm sections. After being dewaxed and rehydrated, the sections were scrubbed with Tris-buffered saline buffer. Then, the sections were incubated with a mixture of TdT and dUTP at 37°C for 120 min following by the slides were treated with 0.3% H_2O_2 in methanol for 15 min. After being washed by PBS, the slides were added by converter-POD at 37°C for 30 min. Following incubation, excess labeling solution is washed off with PBS. 3,3′-Diaminobenzidine (DAB) was used to visualize cell apoptosis, and the DAB color was visualized under the microscope for ∼15 min. Sections were then counterstained with hematoxylin, sealed with neutral gum, and finally examined under a microscope.

Clinical Tissues

A total of 21 patients with LUAD receiving chemotherapy in the Third Xiangya Hospital (Changsha, China) from 2016 to 2018 were included in this study. The inclusion criteria were as follows: (1) histopathological examination confirming LUAD; (2) no indication of using molecular targeted drugs; and (3) no operation, or recurrence after operation, with assessable lesions. The 21 patients included in the study received cisplatin-combined chemotherapy and their sensitivity or resistance to cisplatin was determined by computed tomography (CT) analysis before and after cisplatin treatment. The 21 patients were divided into two groups: the cisplatin-sensitive group ($n = 10$) and the cisplatin-resistant group ($n = 11$). The responses to chemotherapy were scored using a tumor regression grade (TRG) developed by the American Joint Commission on Cancer and the College of American Pathology. We allocated the patients with a TRG of 0 or 1 to the cisplatin-sensitive group and those with TRG 2 or 3 to the cisplatin-sensitive group. The study was approved by the Research Ethics Committee of the Xiangya Third Hospital, and signed informed consent was obtained before each subject participated in the study.

Immunohistochemistry Staining

First, paraffin-embedded tissues were sectioned, dewaxed, hydrated, and antigen-repaired. Next, 50 μL peroxidase-blocking solution and 50 μL non-immune animal serum were added, and the sections were incubated at room temperature for 10 min. The primary antibodies anti-WDHD1 (1:100, ab72436; Abcam) and anti-Ki67 (1:100, GTX16667; Genetex) were added to the sections and incubated overnight at 4°C. Each section, after washing, was incubated at room temperature for 30 min with a drop of biotin-labeled secondary antibody. 3,3′-Diaminobenzidine was used to develop the visual signal. Hematoxylin was used as a counterstain. Two pathologists who were blinded to clinical pathology information scored the samples. The score was determined by the proportion of positive tumor cells and the intensity of staining. Tumor cell proportions were scored as follows: "0" (< 5% positive tumor cells), "1" (5–25% positive tumor cells), "2" (25–50% positive tumor cells), "3" (25–75% positive tumor cells), and "4" (>75% positive tumor cells). Staining intensity was graded according to the following standard: "0" (no staining), "1" (weak staining = light yellow), "2" (moderate staining = yellow brown), and "3" (strong staining = brown). The total immunostaining score (scored as 0, 1, 2, 3, 4, 6, 8, 9, or 12) was calculated as the value of the proportion of positive cells score multiplied by the staining intensity score. Staining index scores ≥6 were identified as high expression, whereas scores <6 were considered low expression.

Statistical Analysis

The data were statistically analyzed using GraphPad Prism 7.0 (GraphPad Inc., California, CA, USA) (21). All experiments were performed in triplicates, and data are expressed as mean ± standard deviation (**Supplementary Material 2**). The significance of the data sets was tested using analysis of variance (ANOVA). Comparison between specific groups was performed using a Student t-test (e.g., quantitative real-time PCR data). Multiple comparisons were performed using Bonferroni test and Tukey test (e.g., flow cytometry, wound healing assay, colony

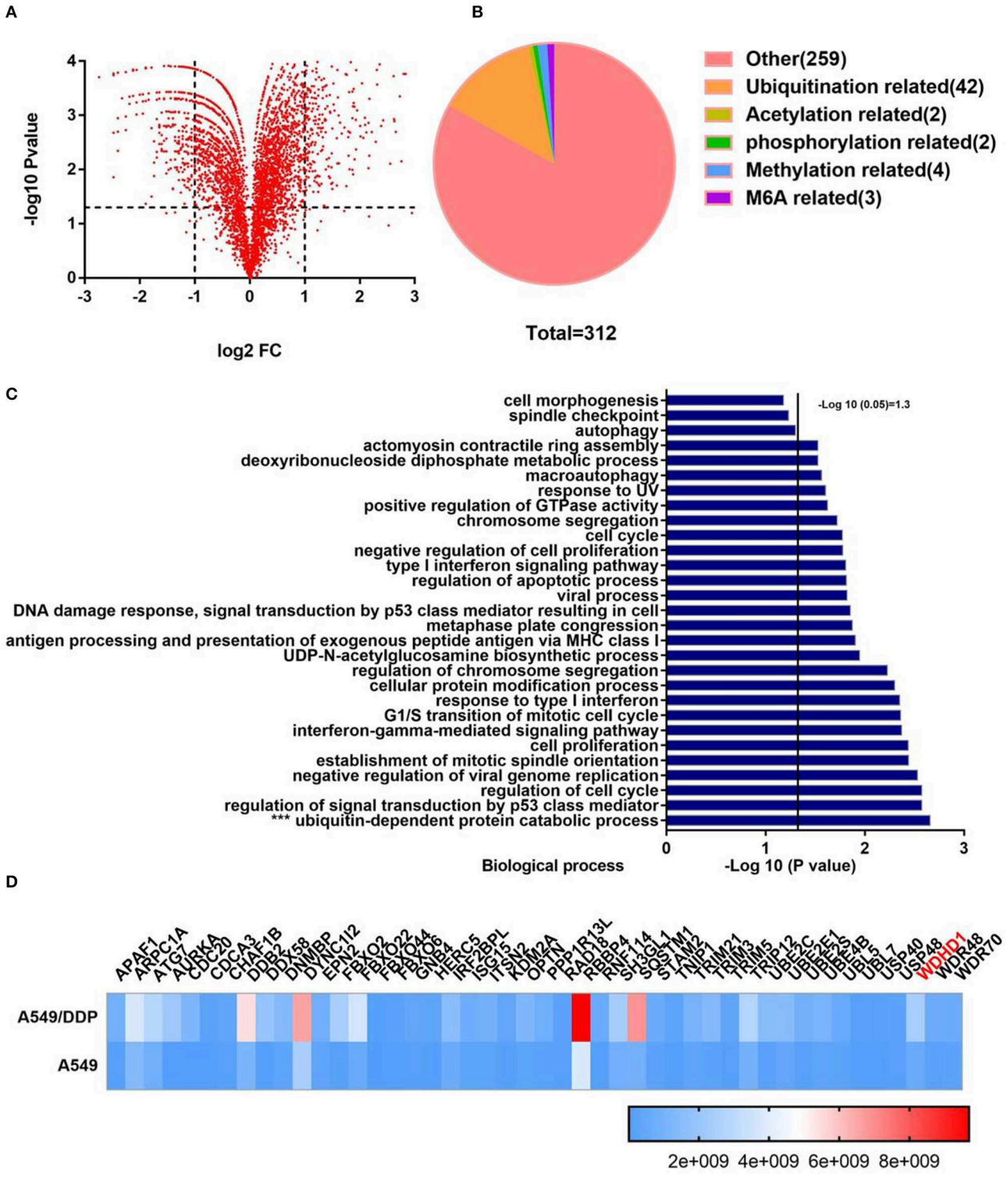


FIGURE 1 | Quantitative analysis of the screened global proteome showed significant change in the protein expression in A549/DDP cells. **(A)** Quantitative analysis of the global proteome identified 5,758 quantitative proteins. Among the quantified proteins in A549/DDP cells, 312 proteins were upregulated, and 345 proteins were downregulated as compared to A549 cells. **(B)** Posttranslational modification-related proteins were significantly upregulated. **(C)** According to an enrichment analysis, ubiquitin-dependent protein catabolism was active. **(D)** The protein abundance and relative expression of 46 ubiquitin-related enzymes in A549/DDP and A549 cells.

formation assay, and MTT assay). $p < 0.05$ in all cases was considered to be statistically significant.

RESULTS

Analysis and Verification of the Global Proteome in A549 and A549/DDP Cells

Ubiquitin-Dependent Protein Catabolism Was More Active in A549/DDP Cells Than That in A549 Cells

Cisplatin resistance in patients with LUAD is one of the main reasons for poor tumor prognosis. To explore the mechanism of cisplatin resistance in LUAD, we performed a quantitative analysis of the global proteome of both A549/DDP and A549 cell lines. A total of 7,475 protein groups were identified, among which 5,758 proteins were quantified. The fold-change cutoff was set when proteins with quantitative ratios >2 or <1/2 were deemed significant ($p < 0.05$). Among the quantified proteins in A549/DDP cells, we found that 312 proteins were upregulated, and 345 proteins were downregulated as compared to A549 cells (**Figure 1A**). Analysis of the significantly upregulated proteins revealed that post-translational modification of related proteins were upregulated (**Figure 1B**). The public database, https://david.ncifcrf.gov/, was used to analyze the functional enrichment of significantly upregulated proteins. We found that ubiquitin-dependent protein catabolism was more active in A549/DDP cells as compared to that in A549 cells, with statistically significant differences ($p < 0.05$) (**Figure 1C**). Ubiquitination-related websites (http://ubibrowser.ncpsb.org/ubibrowser/home/index and http://iuucd.biocuckoo.org/) and protein chip results suggested that 46 ubiquitination-related enzymes were significantly upregulated in A549/DDP cells, with protein abundance as shown in the heat map (**Figure 1D**). These results suggest that protein ubiquitination is associated with cisplatin resistance.

Knockout of the Ubiquitin Ligase WDHD1 Increases Cisplatin Sensitivity in LUAD Cells

Ubiquitin ligases play an important role in recognizing substrates during protein ubiquitination and are closely associated with cisplatin resistance in malignant tumors. Therefore, to understand the role of ubiquitin ligases in cisplatin resistance in LUAD, we consulted the public database http://gepia.cancer-pku.cn/ and found that the upregulation of the ubiquitin ligase WDHD1 (**Figure 2A**) and ARPC1A (**Supplementary Figure 1A**) negatively correlated with patient prognosis. The expression of WDHD1 in the A549/DDP cell line was 2.17 times higher than that in the A549 cell line in a quantitative analysis of the global proteome. Therefore, we speculated that WDHD1 overexpression might be the cause of cisplatin resistance in LUAD. According to the public database http://ualcan.path.uab.edu/index.html, the expression of WDHD1 (**Figure 2B**) and ARPC1A (**Supplementary Figure 1B**) in lung cancer tissues was higher than that in the adjacent tissues, suggesting that the upregulation of WDHD1 and ARPC1A may be one of the reasons for the occurrence and development of LUAD. The level of WDHD1 and ARPC1A expression in A549/DDP and A549 cells was detected by PCR (**Supplementary Figure 1C**). The results showed that the level of WDHD1 expression in A549/DDP cell lines was higher than that in the A549 cell lines ($P < 0.05$), which was consistent with the public database and the trend of the protein chip results (**Figure 2C**). To further explore the effect of WDHD1 on the sensitivity of A549/DDP cells to cisplatin, we constructed siRNA to knock out WDHD1 in A549/DDP cells. MTT assay was used to detect the viability of tumor cells and calculate the IC50 and IC25 for cisplatin toward A549/DDP cells. It was found that the viability of A549/DDP cells significantly decreased after WDHD1 knockout (**Figure 2D**), and the IC50 decreased significantly (**Supplementary Table 1**).

WDHD1 Increases Cisplatin Resistance by Promoting MAPRE2 Ubiquitination in LUAD

WDHD1 Decreases the Expression of MAPRE2 in A549/DDP Cells

Literature suggests that ubiquitin ligases affect the sensitivity of malignant tumor cells to cisplatin primarily via regulating the expression of various substrates. To further explore the specific mechanism of WDHD1-induced cisplatin resistance in LUAD cells, we aimed to identify the downstream target proteins of WDHD1. Because the main function of ubiquitination is to degrade the target proteins after ubiquitination, the downstream target protein should be downregulated and interact with WDHD1. Combining a protein chip and the protein interaction database http://iuucd.biocuckoo.org, 15 proteins were found to be downregulated by 1.5 times or more and interact with WDHD1. Moreover, the ubiquitin ligase target proteins must have ubiquitination sites to be ubiquitinated. By referring to the ubiquitination site prediction database http://www.phosphosite.org, it was suggested that there are seven proteins with ubiquitination sites in the aforementioned downregulated proteins (**Supplementary Table 2**). Combined with the http://ualcan.path.uab.edu/index.html public database, it is suggested that only MAPRE2 and DSTN are expressed at low levels in LUAD (**Supplementary Figure 2A**). Therefore, we speculate that MAPRE2 and DSTN may be the target proteins of WDHD1.

To determine whether WDHD1 can regulate the expression of MAPRE2 and DSTN, the expression of MAPRE2 and DSTN in A549/DDP cells was detected by Western blotting (WB) after knocking out WDHD1 with si-WDHD1. The transfection efficiency of siRNA is shown in **Figure 2E**. The results showed that the knocking out of WDHD1 could increase the expression of MAPRE2 but had no significant effect on DSTN (**Figure 3A** and **Supplementary Figure 2B**). The WB results showed that WDHD1 expression in A549/DDP cells was higher than that in A549 cells (**Figure 3A**). These results were consistent with the protein chip findings, which confirmed the reliability of the chip. To explore whether the downregulation of MAPRE2 can affect the cisplatin sensitivity of LUAD cells, we constructed a siRNA knockout of MAPRE2 (si-MAPRE2) and control si-RNA (si-con). Both A549 and A549/DDP cells were treated with cisplatin

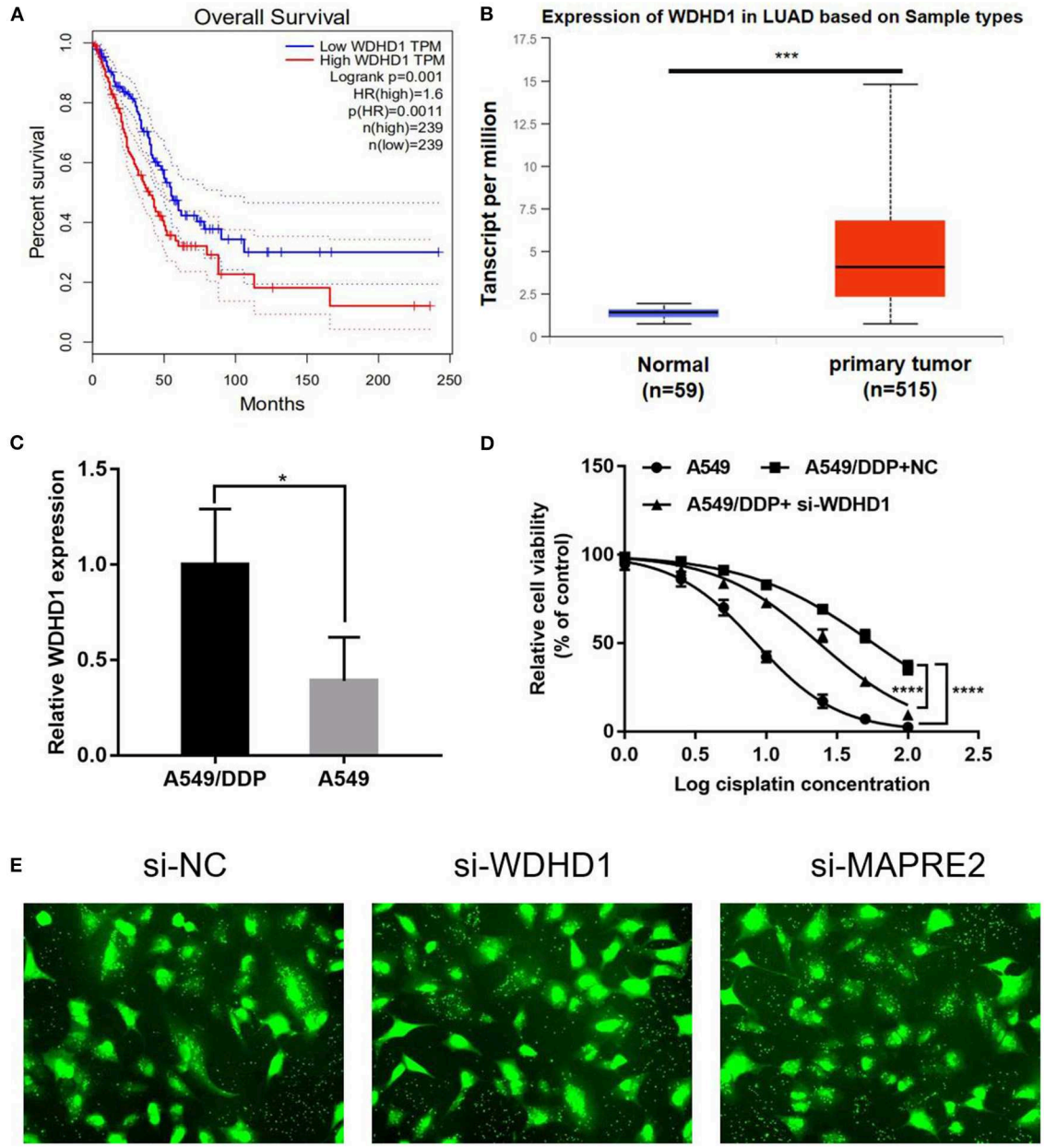


FIGURE 2 | WDHD1 knockout increases cisplatin sensitivity in lung adenocarcinoma cells. **(A)** According to the public database, overexpression of WDHD1 negatively correlated with the prognosis of patients with lung adenocarcinoma. **(B)** Referring to the public database, WDHD1 expression in lung cancer tissues was higher than that in adjacent tissues. **(C)** The PCR results showed that WDHD1 expression in A549/DDP cells was significantly higher than that in A549 cells. **(D)** After WDHD1 was knocked out, the proliferation and IC50 of A549/DDP decreased significantly. **(E)** The transfection efficiency of siRNA. $n = 3$; $^{*}p < 0.05$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

at an IC25 concentration of cisplatin for A549/DDP cells and transfected with Si-Con and si-MAPRE2, respectively. A WB confirmed that MAPRE2 was successfully knocked out, and after MAPRE2 knocked out, A549 cell viability increased (**Figure 3B**). Flow cytometry results indicated that apoptosis was decreased under the action of cisplatin after MAPRE2 was knocked out in A549 and A549/DDP cells (**Figure 3C**). These results suggest that a decrease in MAPRE2 may be the cause of cisplatin resistance in LUAD cells. In conclusion, we speculate that MAPRE2 may act as a downstream target protein of WDHD1 and plays an important role in regulating cellular sensitivity to cisplatin.

WDHD1 Induces Cisplatin Resistance in A549/DDP Cells by Increasing the Ubiquitination Degradation of MAPRE2

Protein molecules perform essential cellular functions by forming protein complexes. To determine whether WDHD1 regulates the expression of MAPRE2 by forming protein complexes, we

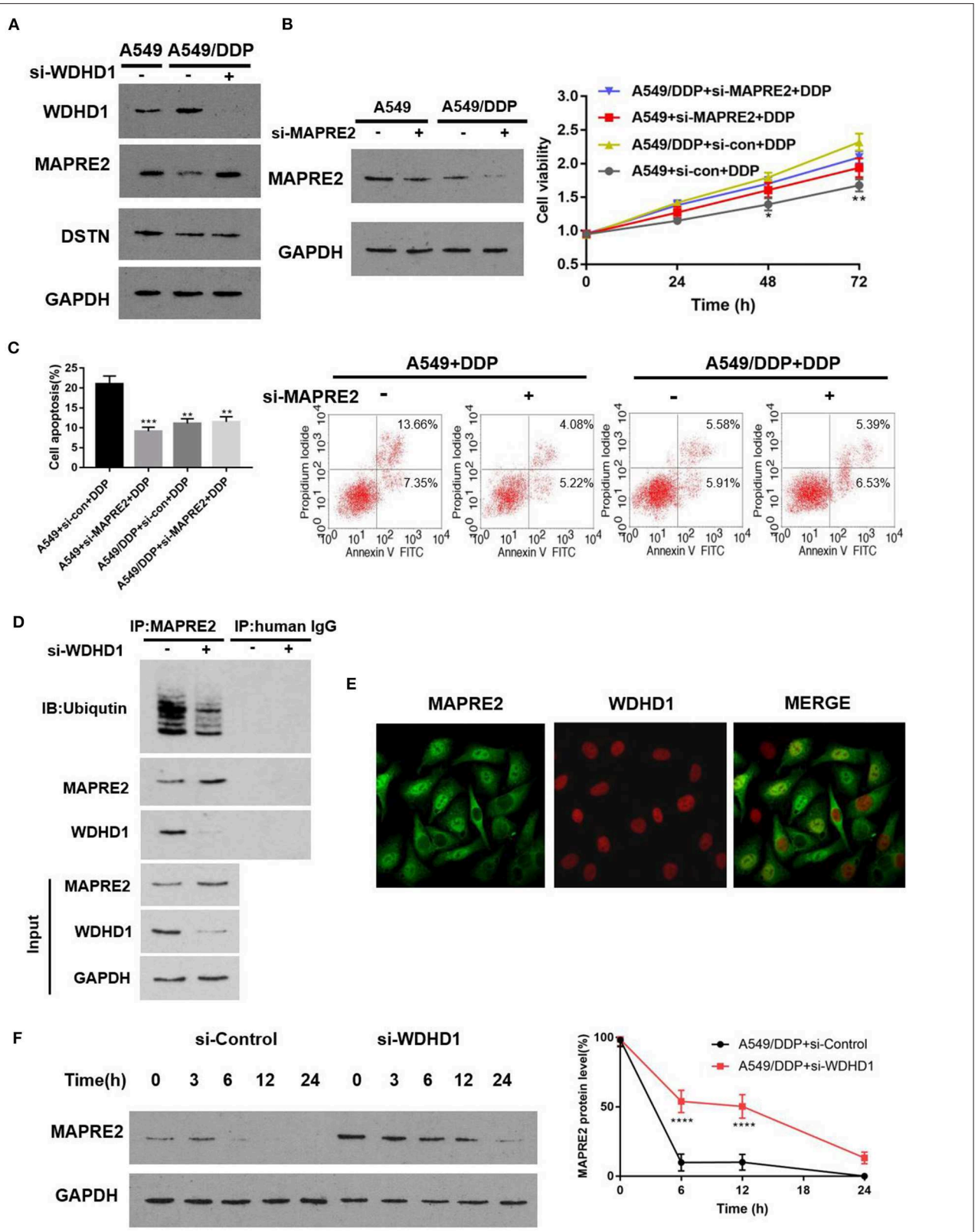


FIGURE 3 | WDHD1 forms a complex with MAPRE2 and increases the ubiquitination degradation of MAPRE2. **(A)** Western blot results indicating that WDHD1 expression was higher in A549/DDP cells than in A549 cells. MAPRE2 was upregulated, and DSTN did not significantly change after knocking out WDHD1. **(B,C)** The

(Continued)

FIGURE 3 | results from an MTT assay and flow cytometry showing that after MAPRE2 was knocked out in A549 cells, the drug sensitivity of lung adenocarcinoma cells to cisplatin decreased, viability increased, and apoptosis decreased. **(D)** Coimmunoprecipitation showing that WDHD1 and MAPRE2 are interacting proteins. **(E)** Immunofluorescence demonstrating that WDHD1 (red) is primarily present in the nucleus, and MAPRE2 (green) is present in both nucleus and cytoplasm. **(F)** Protein stability experiments confirming that MAPRE2 degradation slowed down after knocking out WDHD1. DDP: IC25 concentration of cisplatin for A549/DDP cells. $n = 3$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

investigated the interaction between the ubiquitin ligase WDHD1 and MAPRE2 by coimmunoprecipitation (Co-IP) and detected the level of MAPRE2 ubiquitination. A549/DDP cells transfected with either si-con or si-WDHD1 were established as groups A and B, respectively. Coimmunoprecipitation was performed after successful transfection. The results showed that WDHD1 and MAPRE2 interacted with each other, and the level of MAPRE2 ubiquitination was significantly decreased when WDHD1 was knocked out (**Figure 3D**). The immunofluorescence results showed that WDHD1 (red) primarily existed in the nucleus, and MAPRE2 (green) existed both in the nucleus and cytoplasm, suggesting that the interaction between WDHD1 and MAPRE2 mainly occurred in the nucleus (**Figure 3E**). The protein stability test confirmed the effect of WDHD1 on MAPRE2 degradation. The results showed that the rate of MAPRE2 degradation was significantly reduced when WDHD1 was knocked out (**Figure 3F**). The above experiments confirmed that WDHD1 and MAPRE2 interact with each other, and WDHD1 promotes the degradation of MAPRE2 by ubiquitination.

A functional recovery experiment was conducted to explore the effect of WDHD1 on cisplatin resistance by regulating MAPRE2, by establishing five groups designated A to E. Groups A was transfected with si-control; groups B and C were transfected with si-WDHD1; group D was transfected with both si-WDHD1 and si-MAPRE2; and group E was transfected with si-MAPRE2. Groups C to E were treated with DDP. After the DPP treatment, MTT assay (**Figure 4A**), flow cytometry (**Figure 4B**), scratch test (**Figure 4C**), and cell clone formation experiment (**Figure 4D**) were performed. The results showed that the cellular viability decreased, and apoptosis increased significantly after knocking out WDHD1 in A549/DDP cells. When both WDHD1 and MAPRE2 were knocked out, cellular viability increased, and apoptosis decreased.

Animal experiments were conducted to verify the effect of WDHD1 and MAPRE2 interaction on cisplatin drug sensitivity. In groups A and B, A549/DDP cells were transfected with si-control and si-WDHD1, respectively. Group C was transfected with both si-WDHD1 and si-MAPRE2. After successful transfection, cells in groups A to C were subcutaneously implanted into mice. When the mice developed palpable tumors, they were administered cisplatin. We found that in comparison with group B the growth rate of the tumors in group C was significantly higher (**Figure 5A**). A WB confirmed that WDHD1 expression was upregulated in A549/DDP tumor tissues, and MAPRE2 was upregulated in WDHD1 knockout tumor tissues (**Figure 5B**). The expression of Ki67 in the WDHD1 knockout group was significantly lower than that in the control group (**Figure 5C**). The TUNEL assay results showed that the apoptotic index of the WDHD1 knockout group was significantly increased (**Figure 5D**). Combined with clinical CT and pathological analysis (**Figures 6A,B**), it could be concluded that cisplatin resistance was increased in patients with WDHD1 overexpression. The *in vivo* experiments further demonstrated that WDHD1 knockout increased the sensitivity of LUAD to cisplatin, decreased the proliferation of LUAD cells, and increased apoptosis.

Therefore, cell, animal, and clinical experiments confirmed that upregulation of the ubiquitin ligase WDHD1 can increase the ubiquitination degradation of MAPRE2, leading to cisplatin resistance in LUAD. Thus, WDHD1 may be an important molecular target for the reversal of cisplatin resistance in LUAD patients.

DISCUSSION

In an attempt to explore the specific molecular mechanism of cisplatin resistance in LUAD, this study identified differentially expressed proteins in A549/DDP and A549 cells by a quantitative analysis of the global proteome. The significantly upregulated proteins was enriched by the public database David in A549/DDP compared to A549 cells, and we found that ubiquitin-dependent protein catabolism is active in A549/DDP cells. A total of 42 ubiquitin ligases were identified among the significantly upregulated proteins. Therefore, we hypothesized that abnormal protein ubiquitination may be the cause of cisplatin resistance in LUAD. To date, studies have shown that cisplatin resistance is related to the regulation of ubiquitin ligase. For example, the ubiquitin ligase TRIM65 is highly expressed in non–small cell lung cancer, and cellular experiments have confirmed that it can reduce cisplatin-induced apoptosis by promoting p53 polyubiquitination degradation (22). The ubiquitin ligase TRIM37 has been shown to ubiquitinate adenylate kinase NEMO and activate the nuclear factor κB signaling pathway, which makes esophageal cancer cells cisplatin-resistant. Literature suggests that ubiquitin ligase primarily regulates cisplatin drug sensitivity by recognizing and regulating its substrates (23). Therefore, E3 was the main target of our present study. The public databases suggested that expression of WDHD1 and ARPC1A is increased in lung cancer tissues as compared to adjacent tissues, which negatively correlated with patient prognosis, suggesting that the upregulation of WDHD1 and ARPC1A may be one of the reasons for the occurrence and development of LUAD. Polymerase chain reaction and WB results demonstrated higher expression of WDHD1 in A549/DDP cells than in A549 cells. Moreover, the results of MTT assay revealed that A549/DDP cell viability, and IC50 decreased significantly when WDHD1 was knocked out. Therefore, we speculate that WDHD1 leads to cisplatin resistance in LUAD cells and present a preliminary exploration of the mechanism of this drug resistance.

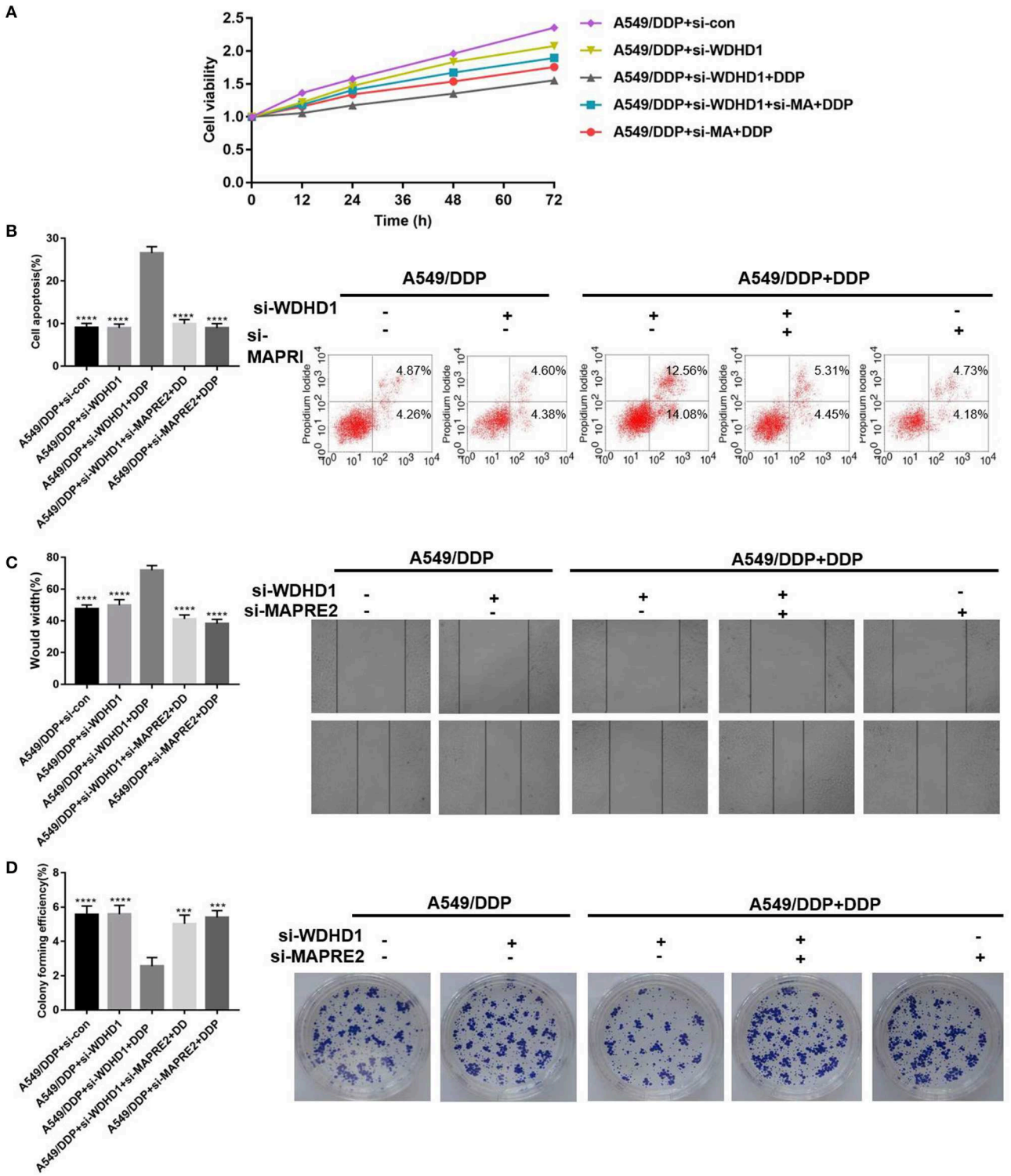


FIGURE 4 | WDHD1 leads to cisplatin resistance by increasing MAPRE2 ubiquitination. **(A)** MTT assay, **(B)** flow cytometry, **(C)** scratch assay, and **(D)** cell colony formation assay results confirming that WDHD1 induced cisplatin resistance in lung adenocarcinoma by downregulating MAPRE2. $n = 3$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

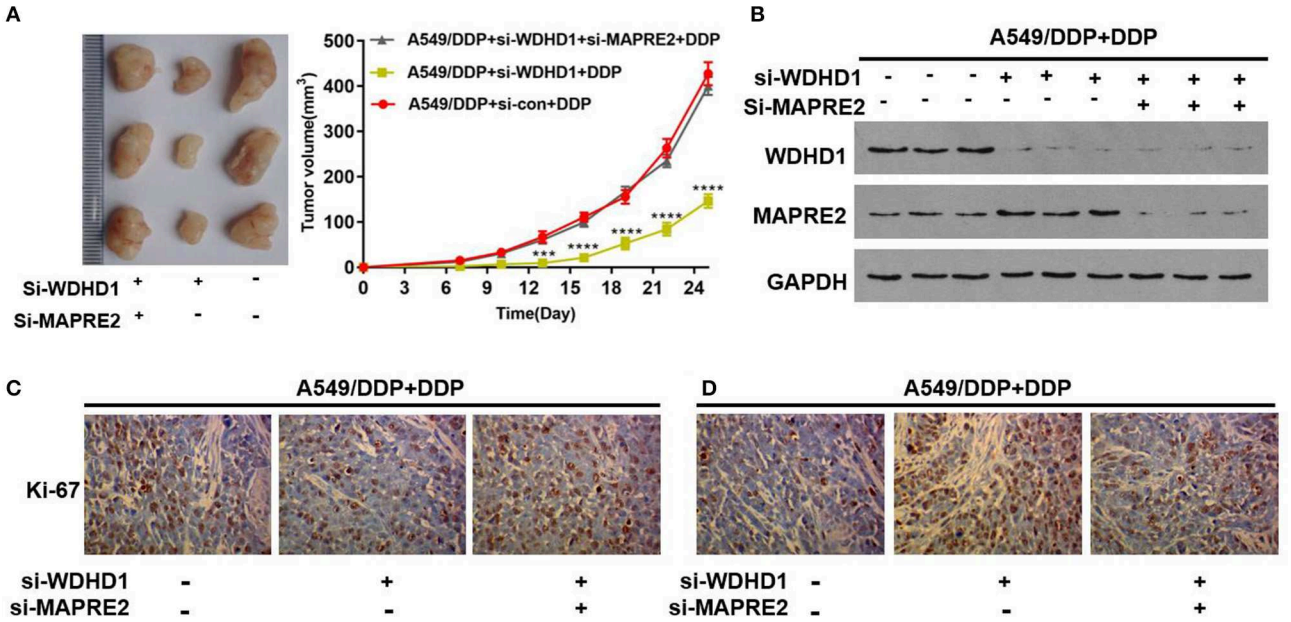


FIGURE 5 | Verification of the function of WDHD1 by animal experiments. **(A)** The transplanted tumor experiment showing that the inhibition of tumor volume by cisplatin in WDHD1 knockout group was more obvious than that in the control group. **(B)** A Western blot indicating that WDHD1 expression was upregulated in mice tissues implanted with A549/DDP cells, and MAPRE2 expression was increased in tumor tissues implanted with WDHD1 knockout cells. **(C,D)** The expression of Ki67 in the A549/DDP WDHD1 knockout group was significantly lower than that in the control group. The apoptotic index detected by a TUNEL assay was significantly higher in A549/DDP WDHD1 knockout group than that in the control group. ***$p < 0.001$; ****$p < 0.0001$. DDP: IC25 concentration for A549/DDP cells.

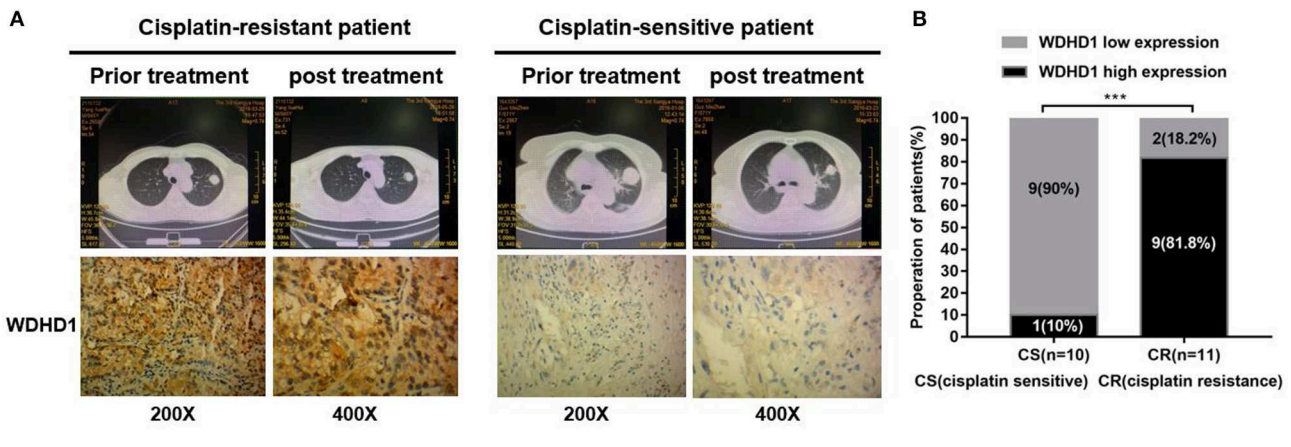


FIGURE 6 | The expression of WDHD1 in cisplatin-resistant lung adenocarcinoma tissues is higher than that in cisplatin-sensitive resistant tissues. **(A)** WDHD1 expression in paraffin-embedded cisplatin-treated lung adenocarcinoma tissues as measured by *in situ* hybridization. Representative cases of cisplatin-sensitive ($n = 10$) and cisplatin-resistant ($n = 11$) patients with low or high WDHD1 expression are shown. **(B)** Proportion of cisplatin-treated lung adenocarcinoma patients with low or high WDHD1 expression. ***$p < 0.001$.

WDHD1, also known as CTF4 or AND-1, is a DNA-binding protein in the nucleoplasm (24) that binds to DNA through the HMG domain (25). Moreover, WDHD1 directly interferes with the formation and progression of replication forks or indirectly interferes with DNA replication by stabilizing DNA helicase complex CMG (26) and interacting with DNA polymerase I (polA) (27), which is required for initiation of DNA replication (28, 29). In addition, studies have shown that WDHD1 regulates the cellular response to DNA damage (30, 31), DNA repair, and mitosis (32). WDHD1 has also been reported to be closely related to malignant tumors. For example, studies have confirmed that WDHD1 is upregulated in cervical cancer cells, and knocking out WDHD1 in these cells leads to G1 stagnation, which affects the cell cycle and cellular replication (33). Similarly, high WDHD1 expression was found to be negatively correlated with non–small cell lung cancer and esophageal squamous cell carcinoma, and

the knockout of WDHD1 could effectively inhibit the growth of both lung and esophageal cancer cells (15); however, no studies have reported the association between WDHD1 and cisplatin sensitivity in LUAD.

On the basis of the quantitative analysis of the global proteome and the public websites, Biocuckoo, Phosphosite, and Ualcan, we speculate that MAPRE2 and DSTN may be the target proteins of WDHD1. In this study, the WB results showed that MAPRE2 was upregulated, and DSTN did not change significantly when WDHD1 was knocked out. Moreover, the MTT assay revealed that the MAPRE2 knockout was associated with increased cell viability and decreased apoptosis in A549 cells, suggesting that decreased MAPRE2 may be the cause of cisplatin resistance in LUAD. MAPRE2 is a microtubule-associated protein involved in microtubule polymerization, which is also essential for spindle symmetry during mitosis (34). Literature reveals that MAPRE2 is associated with the occurrence and development of malignant tumors. It has been reported that MAPRE2 is overexpressed in hepatocellular carcinoma (35) and esophageal squamous cell carcinoma (36); thus, MAPRE2 might be involved in tumorigenesis and promotion of tumor cell growth through Wnt signaling pathway or Aurora-B activation (37, 38). A study by Abiatari et al. (39) demonstrated that overexpression of MAPRE2 is associated with decreased survival and perineural infiltration in pancreatic cancer patients. Therefore, MAPRE2 overexpression in various malignant tumors is positively correlated with tumor growth, nerve infiltration, and poor prognosis. However, using the public database ULACN, it was observed that the levels of MAPRE2 expression in different types of malignant tumors and normal tissues differ from those suggested in the literature. For example, in LUAD, renal papillary carcinoma, bladder urothelial carcinoma, and other cancers, MAPRE2 expression is significantly lower than that in normal tissues. Therefore, MAPRE2 expression may be tissue-specific and plays a differential role in various malignant tumors. According to the protein chip and experimental results, it has been demonstrated that low MAPRE2 expression in LUAD cells may be a contributing factor to cisplatin resistance in LUAD. Moreover, Co-IP results showed that there was interaction between WDHD1 and MAPRE2. Further, immunofluorescence results revealed that WDHD1 interacts with MAPRE2 in the nucleus. Protein stability test results indicated that WDHD1 promotes ubiquitination degradation of MAPRE2. In addition, both functional recovery and animal experiments confirmed that WDHD1 could induce cisplatin resistance by increasing the ubiquitination degradation of MAPRE2. Finally, clinical CT and pathological analysis further confirmed that WDHD1 overexpression increased cisplatin resistance in patients with LUAD.

In conclusion, our results indicate that the ubiquitin ligase WDHD1 induces cisplatin resistance in LUAD by promoting the degradation of MAPRE2. These findings suggest that WDHD1 and MAPRE2 may be potential biomarkers of cisplatin sensitivity in patients with LUAD. Although the regulatory mechanism of ubiquitin ligase in tumorigenesis is highly complex, our results indicate that the level of ubiquitin ligase expression and its substrates has the potential to predict cisplatin sensitivity in LUAD. Although cellular, animal, and clinical data were presented in this study, something remain poorly understood, for example, effect of overexpression of MAPRE2 in A549 on cisplatin sensitivity, specific mechanism of MAPRE2 regulating cisplatin drug sensitivity, and the specific sites and types of ubiquitination. To explore the specific mechanism of MAPRE2 regulating cisplatin drug sensitivity, we use the public database https://string-db.org/ to explore MAPRE2′ interaction protein (**Supplementary Figure 3A**) and search their expression level and change through our protein chip (**Supplementary Figure 3B**). Among them, AURKB is significantly high expression in A549/DDP, and it has been confirmed that its high expression is related to cisplatin resistance (40, 41). Then, DNA damage repair gene set was obtained from GSEA, and the correlation between MAPRE2 and the DNA damage repair gene was analyzed by String. Two proteins were related to MAPRE2 (**Supplementary Figure 3C**). RAE1 was highly expressed in A549/DDP, and its high expression was negatively related to the prognosis of LUAD patients (**Supplementary Figure 3D**). MAPRE2 may affect cisplatin drug sensitivity by regulating AURKB and DNA damage repair–related proteins. Thus, we will overexpress MAPRE2 in A549/DDP to further confirm its effect on cisplatin drug sensitivity and explore the specific mechanism of MAPRE2 affecting cisplatin drug sensitivity from the perspective of DNA damage repair and AURKB regulation in our further experiment. In addition, future studies should focus on understanding the specific sites and types of ubiquitination. Moreover, further exploration of the relationship between ubiquitin ligase and cisplatin sensitivity in LUAD will promote a better understanding of tumor biology. It is important to consider combining the basic factors of cancer biology with clinical practice, which will provide new ideas for the clinical reversal of drug resistance in LUAD, and promote the development of effective treatment strategies.

In summary, this study demonstrates that the ubiquitin ligase WDHD1 is overexpressed in LUAD and plays an important role in cisplatin resistance by promoting MAPRE2 ubiquitination. Our findings indicated a novel molecular mechanism underlying cisplatin resistance in LUAD. Thus, WDHD1 and MAPRE2 could serve as novel therapeutic targets for reversal of cisplatin resistance in LUAD.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Third Xiangya Hospital of Central South University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Ethics Committee for the Animal Laboratory of Central South University. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

LG and KC conceived the study and performed data interpretation. LG and MX performed the experiments, analyzed data, and wrote the manuscript. DH, YH, YZhu, LX, YB, XL, QZ, JL, MZ, YZho, YC, YZha, LD, RZ, and HL participated in the experiments.

FUNDING

This work was supported by the National Natural Science Foundation of China (81874137), the Outstanding Youth Foundation of Hunan Province (2018JJ1047), the Huxiang Young Talent Project (2016RS3022), Hunan Province Science and technology talent promotion project. (2019TJ-Q10), and Project of scientific research plan of health and Family Planning Commission of Hunan Province (c20180476). The Independent Exploration and Innovation Project of Central South University (2018zzts938).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.00461/full#supplementary-material

REFERENCES

1. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* (2016) 66:115–32. doi: 10.3322/caac.21338
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
3. Beretta GL, Gatti L, Tinelli S, Corna E, Colangelo D, Zunino F, et al. Cellular pharmacology of cisplatin in relation to the expression of human copper transporter CTR1 in different pairs of cisplatin-sensitive and -resistant cells. *Biochem Pharmacol.* (2004) 68:283–91. doi: 10.1016/j.bcp.2004.03.022
4. Sawant A, Kothandapani A, Zhitkovich A, Sobol RW, Patrick SM. Role of mismatch repair proteins in the processing of cisplatin interstrand cross-links. *DNA Repair (Amst).* (2015) 35:126–36. doi: 10.1016/j.dnarep.2015.10.003
5. Zhu H, Luo H, Zhang W, Shen Z, Hu X, Zhu X. Molecular mechanisms of cisplatin resistance in cervical cancer. *Drug Des Devel Ther.* (2016) 10:1885–95. doi: 10.2147/DDDT.S106412
6. Fang C, Chen YX, Wu NY, Yin JY, Li XP, Huang HS, et al. MiR-488 inhibits proliferation and cisplatin sensibility in non-small-cell lung cancer (NSCLC) cells by activating the eIF3a-mediated NER signaling pathway. *Sci Rep.* (2017) 7:40384. doi: 10.1038/srep40384
7. Feng X, Liu H, Zhang Z, Gu Y, Qiu H, He Z. Annexin A2 contributes to cisplatin resistance by activation of JNK-p53 pathway in non-small cell lung cancer cells. *J Exp Clin Cancer Res.* (2017) 36:123. doi: 10.1186/s13046-017-0594-1
8. Lai YH, Kuo C, Kuo MT, Chen HHW. Modulating chemosensitivity of tumors to platinum-based antitumor drugs by transcriptional regulation of copper homeostasis. *Int J Mol Sci.* (2018) 19:E1486. doi: 10.3390/ijms19051486
9. Matsuura K, Huang NJ, Cocce K, Zhang L, Kornbluth S. Downregulation of the proapoptotic protein MOAP-1 by the UBR5 ubiquitin ligase and its role in ovarian cancer resistance to cisplatin. *Oncogene.* (2017) 36:1698–706. doi: 10.1038/onc.2016.336
10. Hua Z, Vierstra RD. The cullin-RING ubiquitin-protein ligases. *Annu Rev Plant Biol.* (2011) 62:299–334. doi: 10.1146/annurev-arplant-042809-112256
11. Walter RFH, Werner R, Wessolly M, Mairinger E, Borchert S, Schmeller J, et al. Inhibition of MDM2 via Nutlin-3A: a potential therapeutic approach for pleural mesotheliomas with MDM2-induced inactivation of wild-type P53. *J Oncol.* (2018) 2018:1986982. doi: 10.1155/2018/1986982
12. Zhu J, Zhao C, Zhuang T, Jonsson P, Sinha I, Williams C, et al. RING finger protein 31 promotes p53 degradation in breast cancer cells. *Oncogene.* (2016) 35:1955–64. doi: 10.1038/onc.2015.260
13. Song YH, Zhang CQ, Chen FF, Lin XY. Upregulation of neural precursor cell expressed developmentally downregulated 4–1 is associated with poor prognosis and chemoresistance in lung adenocarcinoma. *Chin Med J (Engl).* (2018) 131:16–24. doi: 10.4103/0366-6999.221262
14. Liu Z, Wu Y, Tao Z, Ma L. E3 ubiquitin ligase Hakai regulates cell growth and invasion, and increases the chemosensitivity to cisplatin in nonsmallcell lung cancer cells. *Int J Mol Med.* (2018) 42:1145–51. doi: 10.3892/ijmm.2018.3683
15. Sato N, Koinuma J, Fujita M, Hosokawa M, Ito T, Tsuchiya E, et al. Activation of WD repeat and high-mobility group box DNA binding protein 1 in pulmonary and esophageal carcinogenesis. *Clin Cancer Res.* (2010) 16:226–39. doi: 10.1158/1078-0432.CCR-09-1405
16. Laurila E, Savinainen K, Kuuselo R, Karhu R, Kallioniemi A. Characterization of the 7q21-q22 amplicon identifies ARPC1A, a subunit of the Arp2/3 complex, as a regulator of cell migration and invasion in pancreatic cancer. *Genes Chromosomes Cancer.* (2009) 48:330–9. doi: 10.1002/gcc.20643
17. Liu JY, Zeng QH, Cao PG, Xie D, Chen X, Yang F, et al. RIPK4 promotes bladder urothelial carcinoma cell aggressiveness by upregulating VEGF-A through the NF-kappaB pathway. *Br J Cancer.* (2018) 118:1617–27. doi: 10.1038/s41416-018-0116-8
18. Yu L, Sun Y, Li J, Wang Y, Zhu Y, Shi Y, et al. Silencing the Girdin gene enhances radio-sensitivity of hepatocellular carcinoma via suppression of glycolytic metabolism. *J Exp Clin Cancer Res.* (2017) 36:110. doi: 10.1186/s13046-017-0580-7
19. Jia X, Zhou H, Wu C, Wu Q, Ma S, Wei C, et al. The ubiquitin ligase RNF125 targets innate immune adaptor protein TRIM14 for ubiquitination and degradation. *J Immunol.* (2017) 198:4652–8. doi: 10.4049/jimmunol.1601322
20. Zeng Q, Liu J, Cao P, Li J, Liu X, Fan X, et al. Inhibition of REDD1 sensitizes bladder urothelial carcinoma to paclitaxel by inhibiting autophagy. *Clin Cancer Res.* (2018) 24:445–59. doi: 10.1158/1078-0432.CCR-17-0419
21. Zhu Y, He D, Bo H, Liu Z, Xiao M, Xiang L, et al. The MRVI1-AS1/ATF3 signaling loop sensitizes nasopharyngeal cancer cells to paclitaxel by regulating the Hippo-TAZ pathway. *Oncogene.* (2019) 38:6065–81. doi: 10.1038/s41388-019-0858-7
22. Li Y, Ma C, Zhou T, Liu Y, Sun L, Yu Z. TRIM65 negatively regulates p53 through ubiquitination. *Biochem Biophys Res Commun.* (2016) 473:278–82. doi: 10.1016/j.bbrc.2016.03.093
23. Wu G, Song L, Zhu J, Hu Y, Cao L, Tan Z, et al. An ATM/TRIM37/NEMO axis counteracts genotoxicity by activating nuclear-to-cytoplasmic NF-κB signaling. *Cancer Res.* (2018) 78:6399–412. doi: 10.1158/0008-5472.CAN-18-2063
24. Kohler A, Schmidt-Zachmann MS, Franke WW. AND-1, a natural chimeric DNA-binding protein, combines an HMG-box with regulatory WD-repeats. *J Cell Sci.* (1997) 110(Pt 9):1051–62.
25. Guan C, Li J, Sun D, Liu Y, Liang H. The structure and polymerase-recognition mechanism of the crucial adaptor protein AND-1 in the human replisome. *J Biol Chem.* (2017) 292:9627–36. doi: 10.1074/jbc.M116.758524
26. Hao J, de Renty C, Li Y, Xiao H, Kemp MG, Han Z, et al. And-1 coordinates with Claspin for efficient Chk1 activation in response to replication stress. *EMBO J.* (2015) 34:2096–110. doi: 10.15252/embj.201488016
27. Moiseeva T, Hood B, Schamus S, O'Connor MJ, Conrads TP, Bakkenist CJ. ATR kinase inhibition induces unscheduled origin firing through a Cdc7-dependent association between GINS and And-1. *Nat Commun.* (2017) 8:1392. doi: 10.1038/s41467-017-01401-x
28. Im JS, Ki SH, Farina A, Jung DS, Hurwitz J, Lee JK. Assembly of the Cdc45-Mcm2–7-GINS complex in human cells requires the Ctf4/And-1,

RecQL4, and Mcm10 proteins. *Proc Natl Acad Sci USA*. (2009) 106:15628–32. doi: 10.1073/pnas.0908039106
29. Abe T, Kawasumi R, Giannattasio M, Dusi S, Yoshimoto Y, Miyata K, et al. AND-1 fork protection function prevents fork resection and is essential for proliferation. *Nat Commun.* (2018) 9:3091. doi: 10.1038/s41467-018-05586-7
30. Chen Y, Liu H, Zhang H, Sun C, Hu Z, Tian Q, et al. And-1 coordinates with CtIP for efficient homologous recombination and DNA damage checkpoint maintenance. *Nucleic Acids Res.* (2017) 45:2516–30. doi: 10.1093/nar/gkw1212
31. Li Y, Li Z, Wu R, Han Z, Zhu W. And-1 is required for homologous recombination repair by regulating DNA end resection. *Nucleic Acids Res.* (2017) 45:2531–45. doi: 10.1093/nar/gkw1241
32. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform.* (2011) 12:449–62. doi: 10.1093/bib/bbr042
33. Hsieh CL, Lin CL, Liu H, Chang YJ, Shih CJ, Zhong CZ, et al. WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. *Nucleic Acids Res.* (2011) 39:4048–62. doi: 10.1093/nar/gkq1338
34. Bruning-Richardson A, Langford KJ, Ruane P, Lee T, Askham JM, Morrison EE. EB1 is required for spindle symmetry in mammalian mitosis. *PLoS ONE.* (2011) 6:e28884. doi: 10.1371/journal.pone.0028884
35. Fujii K, Kondo T, Yokoo H, Yamada T, Iwatsuki K, Hirohashi S. Proteomic study of human hepatocellular carcinoma using two-dimensional difference gel electrophoresis with saturation cysteine dye. *Proteomics.* (2005) 5:1411–22. doi: 10.1002/pmic.200401004
36. Wang Y, Zhou X, Zhu H, Liu S, Zhou C, Zhang G, et al. Overexpression of EB1 in human esophageal squamous cell carcinoma (ESCC) may promote cellular growth by activating beta-catenin/TCF pathway. *Oncogene.* (2005) 24:6637–45. doi: 10.1038/sj.onc.1208819
37. Liu M, Yang S, Wang Y, Zhu H, Yan S, Zhang W, et al. EB1 acts as an oncogene via activating beta-catenin/TCF pathway to promote cellular growth and inhibit apoptosis. *Mol Carcinog.* (2009) 48:212–9. doi: 10.1002/mc.20471
38. Dong X, Liu F, Sun L, Liu M, Li D, Su D, et al. Oncogenic function of microtubule end-binding protein 1 in breast cancer. *J Pathol.* (2010) 220:361–9. doi: 10.1002/path.2662
39. Abiatari I, Gillen S, DeOliveira T, Klose T, Bo K, Giese NA, et al. The microtubule-associated protein MAPRE2 is involved in perineural invasion of pancreatic cancer cells. *Int J Oncol.* (2009) 35:1111–6. doi: 10.3892/ijo_00000426
40. Han X, Zhang JJ, Han ZQ, Zhang HB, Wang ZA. Let-7b attenuates cisplatin resistance and tumor growth in gastric cancer by targeting AURKB. *Cancer Gene Ther.* (2018) 25:300–8. doi: 10.1038/s41417-018-0048-8
41. Yu J, Zhou J, Xu F, Bai W, Zhang W. High expression of Aurora-B is correlated with poor prognosis and drug resistance in non-small cell lung cancer. *Int J Biol Markers.* (2018) 33:215–21. doi: 10.1177/1724600817753098

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 © 2020 Gong, Xiao, He, Hu, Zhu, Xiang, Bao, Liu, Zeng, Liu, Zhou, Zhou, Cheng, Zhang, Deng, Zhu, Lan and Cao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) 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 Oncology

ORIGINAL RESEARCH
published: 16 June 2020
doi: 10.3389/fonc.2020.00963

Proteomic and Phosphoproteomic Maps of Lung Squamous Cell Carcinoma From Chinese Patients

Lulu Pan[1,2†], Xijun Wang[3†], Longhai Yang[4,5†], Lei Zhao[1], Linhui Zhai[1], Junyu Xu[1], Yikun Yang[4], Yousheng Mao[4], Shujun Cheng[3], Ting Xiao[3] and Minjia Tan[1,2*]*

[1] Chemical Proteomics Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China, [2] School of Pharmacy, University of Chinese Academy of Sciences, Beijing, China, [3] State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [4] Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [5] Department of Cardiothoracic Surgery, Shenzhen University General Hospital/Shenzhen University Clinical Medical Academy, Shenzhen, China

OPEN ACCESS

Edited by:
Bhaswati Chatterjee, National Institute of Pharmaceutical Education and Research, India

Reviewed by:
Marco Sciacovelli, University of Cambridge, United Kingdom
Daniele Vergara, University of Salento, Italy

***Correspondence:**
Ting Xiao
xiaot@cicams.ac.cn
Minjia Tan
mjtan@simm.ac.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *31 March 2020*
Accepted: *15 May 2020*
Published: *16 June 2020*

Citation:
Pan L, Wang X, Yang L, Zhao L, Zhai L, Xu J, Yang Y, Mao Y, Cheng S, Xiao T and Tan M (2020) Proteomic and Phosphoproteomic Maps of Lung Squamous Cell Carcinoma From Chinese Patients. Front. Oncol. 10:963. doi: 10.3389/fonc.2020.00963

Lung squamous cell carcinoma (LUSC) is one of the leading causes of tumor-driven deaths in the world. To date, studies on the tumor heterogeneity of LUSC at genomic level have only revealed limited therapeutic benefits. Therefore, system-wide research of LUSC at proteomic level may further improve precision medicine strategies on individual demands. To this end, we performed proteomic and phosphoproteomic study for LUSC samples of 25 Chinese patients. From our results, two subgroups (Cluster I and II) based on proteomic data were identified, which were associated with distinct molecular characteristics and clinicopathologic features. Combined with phosphoproteomic data, our result showed that spliceosome pathway was enriched in Cluster I, while focal adhesion pathway, immune-related pathways and Ras signaling pathway were enriched in Cluster II. In addition, we found that lymph node metastasis (LNM) was associated with our proteomic subgroups and cell cycle pathway was enriched in patients with LNM. Further analysis showed that MCM2, a DNA replication licensing factor involved in cell cycle pathway, was highly expressed in patients with poor prognosis, which was further proved by immunohistochemistry (IHC) analysis. In summary, our study provided a resource of the proteomic and phosphoproteomic features of LUSC in Chinese patients.

Keywords: lung squamous cell carcinoma, proteomics, phosphoproteomics, clustering, lymph node metastasis, cell cycle

INTRODUCTION

Lung cancer is the most malignant tumor with the highest morbidity and mortality in the world (1, 2). Non-small cell lung cancer (NSCLC), as the most common histological types, accounts for more than 80% of all types of lung cancer, in which lung adenocarcinoma and lung squamous cell carcinoma (LUSC) are two major histopathological subtypes (3). In previous studies, LUSC was identified with multiple mutations in cancer driver genes such as *TP53* and *PTEN* (4), and four LUSC mRNA expression subtypes (primitive, classical, secretory, and basal) related with different biological processes (proliferation, xenobiotic metabolism, immune response, cell adhesion) were identified (5).

However, there is still a lack of effective targeted therapies, except a few immunotherapies targeting at PD1 and PD-L1 (6–8). In contrast to genetic features, proteomic characteristics are more directive to reflect the pattern of LUSC as proteins are the "executioners of life" (9, 10).

Recently, a proteogenomic study on LUSC from Western patients has been conducted, which identified three proteomic subtypes associated with immune biology (inflammatory cluster), oxidation-reduction biology (redox cluster) and biology associated with Wnt/stromal signaling (mixed cluster). This study provided a resource and suggested therapeutic strategies based on metabolism and immune for LUSC in Western countries (11).

Global cancer statistics indicated that the occurrence of lung cancer has been decreasing in Western countries but increasing in developing countries (12, 13). In China, lung cancer ranks the first among all malignant tumors due to its high incidence and mortality rates (14). Preliminary researches indicated distinct genomic features of lung cancer for Chinese patients (15). For example, in NSCLC, *EGFR* mutation rate and *EGFR* mutational signatures associated with the inflammatory microenvironments were significantly higher in Chinese patients than those in Western patients. In terms of LUSC, Chinese patients had more frameshift indels in *CDKN2A* and more mutations in *NFE2L2*. Therefore, proteomic studies of LUSC on Western patients may not completely unveil the molecular features of LUSC at proteomic level from Chinese patients.

A previous proteomic study of 10 Chinese patients investigated the possible mechanism of bronchial epithelial carcinogenesis and identified several molecules for early detection, such as GSTP1, HSP1B, and CKB (16). However, systematic proteomic study of LUSC sample in large Chinese cohort is still limited. Besides, this previous work also did not investigate protein phosphorylation, an important protein post-translational modification essential for signaling conduction in cancer (17–23). Exploring molecular mechanisms on proteome and phosphoproteome of LUSC in a Chinese cohort will provide valuable information for the development of targeted therapy.

In this study, the proteomic and phosphoproteomic characteristics of LUSC samples in China were explored. Two subtypes based on proteomic and phosphoproteomic features were acquired. Combining clinicopathologic features, we unveiled that lymph node metastasis (LNM) was associated with clustering and related with patient prognosis possibly through cell cycle pathway.

MATERIALS AND METHODS

Sample Collection

All biospecimens were obtained from the Cancer Institute/Hospital, Peking Union Medical College & Chinese Academy of Medical Sciences and Beijing Xuanwu Hospital, with the approval of the Research Ethics Committee at these two hospitals. By postoperative pathological analysis, these biospecimens were diagnosed as LUSC without other malignant tumors. All patients did not receive any radiotherapy, chemotherapy intervention or targeted therapy before surgery. The postoperative biospecimens were washed with physiological saline on ice to remove blood, and then directly frozen in liquid nitrogen for proteomic research. Judging tumor purity based on HE-stained slides, 25 LUSC tissues with tumor purity>50% were selected to build proteomic profiling and phosphoproteomic profiling. The clinical information was shown in **Supplementary Table S1**.

Protein Extraction

Protein extraction was proceeded in refrigeration room to avoid protein digestion. The tissues were first washed by phosphate-buffered saline (PBS) and quickly dissected using surgical scissors. Then the total protein was extracted by 8 M Urea in 100 mM NH_4HCO_3 (pH 8.0) containing Protease Inhibitor (Roche) and Phosphatase Inhibitor (Roche) on ice for 30 min, followed by 3 min sonication under the condition of 3 s on and 5 s off with 30% power of JY92-IIN (NingBoXinZhi, China). Finally, the protein solution was collected after centrifugation, and the concentration was measured by BCA protein quantification kit (Beyotime Biotechnology, China).

In-solution Digestion

Reaction of reduction was conducted at 56°C for 30 min, which was followed by alkylation reaction at 25°C for 30 min without light. Then the total protein was subjected to LysC (Mass Spec Grade, Hualishi Scientific) with a protease: protein ratio of 1:100 (w/w) for digestion at 37°C for 3 h. After being diluted by four times with 100 mM NH_4HCO_3 (pH 8.0), the protein was digested with trypsin (Mass Spec Grade, Hualishi Scientific) with a protease: protein ratio of 1:50 (w/w) at 37°C for 16 h. The digested peptides were then desalted by SepPak C18 cartridges (Waters, Milford, MA).

TMT 6-Plex Labeling

Internal reference was adopted in the TMT labeling experiment. For the internal reference, 6 tumor samples including different clinical stages as well as differentiation were selected, the peptides were mixed in equal amount as the internal reference in each batch of TMT labeling experiments.

In TMT 6-plex labeling experiment, 200 μg peptides were labeled in every channel of each batch. The "internal reference" peptides were labeled with channel 126, and five tumor peptides were labeled with other channels 127–131, respectively. The isobaric labeling experiment was performed under the instruction of TMT kit. Briefly, TMT reagents were dissolved in ACN and added into peptides in 100 mM triethylammonium bicarbonate (TEAB). The labeling reaction was incubated for 1 h at room temperature, then the labeling reaction was quenched by 5% hydroxylamine for 15 min. After labeling efficiency test (the percentage of TMT modification at lysine residue and peptide N-termini >95%), the labeled peptides were combined at equal amounts and then desalted by SepPak C18 cartridges (Waters, Milford, MA). The 25 tumor samples were finally labeled into five batches in the TMT 6-plex experiment.

HPLC Fractionation for Proteomic Analysis

TMT-labeled peptides (200 μg) in every batch were fractionated by reverse phase XBridge Prep C18 column (250×4.6 mm column containing 5 μm particles, Waters) using an Agilent 1100 HPLC System through a gradient from 3 to 90% buffer B (buffer A: 2% ACN, 98% H_2O; buffer B: 98 % ACN, 2% H_2O, the pH of both buffer A and buffer B was adjusted by NH_4OH to pH = 10.0) with a flow rate of 1 ml/min for 90 min. Twenty fractions were obtained for every batch.

TiO_2 Enrichment for Phosphoproteomic Analysis

TMT-labeled peptides (1 mg) in each batch were subjected to enrichment by titanium dioxide (TiO_2) as previously described (24). Briefly, TMT-labeled peptides were incubated with TiO_2 beads (GL Science, Japan) in loading buffer (5% TFA, 70% ACN, 1 M lactic acid) at 25°C for 30 min. Peptides with nonspecific binding were washed away from beads with washing buffer (0.5% TFA, 70% ACN). Enriched phosphorylated peptides were eventually eluted with 4% ammonium hydroxide and fractionated by home-made C18 tip (3 μm particle size, Agela Technologies Inc.) into 6 fractions.

LC-MS/MS Analysis

Proteomic and phosphoproteomic fractions were analyzed by Orbitrap Fusion following an EASY-nLC 1000 system (Thermo Fisher Scientific). A homemade reverse-phase C18 column (20 cm × 75 μm column containing 3 μm particle, Dikma Technologies Inc.) was used to separate peptides further through a gradient from 5 to 90% buffer B (buffer A: 0.1% FA in 2% ACN, buffer B: 0.1% FA in 90% ACN) in 70 min for proteomic analysis and in 110 min for phosphoproteomic analysis, respectively. Following nanoflow HPLC, Orbitrap precursor spectra were collected from m/z 450–1,500 (proteomic analysis) and m/z 350–1,500 (phosphoproteomic analysis) with a resolution of 60,000 at m/z 200, AGC of 5.0e5 and maximum injection time of 50 ms. In MS/MS acquisition, the top 15 precursors (proteomic analysis) and top 20 precursors (phosphoproteomic analysis) with intensity above 50,000 were selected to be fragmentized by Higher-energy Collision Dissociation with the normalized collision energy of 40%, then the fragment ions were detected in the Orbitrap with a resolution of 15,000 at m/z 200, the isolation window of 1 m/z, AGC of 5.0e5, dynamic exclusion of 50 s and maximum injection time of 80 ms.

Database Searching

Raw files were processed by Proteome Discoverer (PD, version 2.2.0.388; Thermo Fisher Scientific) with the SEQUEST HT search engine against the UniProt human protein database (06/12/2018, 95,549 sequences) (25). TMT 6-plex was chosen as a method for quantification. Acetylation (+42.0105 Da) on protein N-termini and oxidation (+15.9949 Da) on Methionine (M) were designated as dynamic modifications. TMT 6-plex (+229.1629 Da) on Lysine (K) and peptide N-termini, and carbamidomethyl (+57.0215 Da) on Cysteine (C) were set as a static modification. For phosphoproteomic analysis, variable modifications also included phosphorylation (+79.9663 Da) on serine/threonine/tyrosine (S/T/Y). Trypsin/P was set as a specific enzyme with no more than two missed cleavages. The tolerances of MS and MS/MS were set at 10 ppm and 0.02 Da, respectively. The Percolator algorithm (26) in PD was adopted to control peptide spectrum matches at a false discovery rate (FDR) <1% and maximum delta Cn = 0.05. The cutoff of FDR at protein level was set as 1%. For identification of phosphosites, the localization probability threshold was set as 75% which was calculated by the ptmRS algorithm (27).

Proteomic and Phosphoproteomic Data Analysis

Sample Quality Control and Data Normalization

Protein or phosphopeptide intensity was normalized by the median in each channel of five batches of TMT 6-plex experiments through total proteins to calibrate sample loading differences. The phosphosite and phosphoprotein intensity were derived from the sum of phosphopeptide and phosphosite intensity, respectively. For batch assessment, QC of internal reference was analyzed by Pearson correlation. To rule out abnormal samples, unimodal (Gaussian or normal) distribution was tested by a density plot of $\log_2$-transformed TMT ratios for the proteins. For data normalization, $\log_2$ TMT ratios for the proteins or phosphosites /phosphoprotein were normalized by z-score in each sample.

Proteomic and Phosphoproteomic Clustering

Robust proteomic clusters were derived by consensus clustering (28), using the proteins with (1) no missing values; and (2) the top 1,000 most varied proteins within twenty-four tumors. Robust phosphoproteomic clusters were derived by consensus clustering, using the phosphoproteins with (1) the number of missing values<20%; and (2) the top 1,000 most varied proteins within twenty-four tumors. For phosphoproteins, the missing values were imputated using a KNN algorithm. The data sets were clustered by k-means with k from 2 to 8 using the ConsensusClusterPlus R package. The consensus matrix, consensus CDF, delta area plot and silhouette plots were used to assess the appearance of different k values.

Pathway Enrichment Analysis

For proteomic data, gene set enrichment analysis (GSEA) (29) was conducted using gene set database "c2.cp.kegg.v6.2.symbols.gmt" from the MSigDB. For phosphoproteomic data, DAVID bioinformatics tool (30) was used to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (31).

Immunohistochemistry (IHC) and Scoring

Paraffin-embedded LUSC tissue microarray (TMA) was purchased from Shanghai Outdo Biotechnology Company (Shanghai, China), which included 75 cases of LUSC patients with complete clinical pathology data and follow-up information. The TMA sections were baked at 65°C for 4 h and deparaffinized by xylene and ethanol, then incubated with 3% H_2O_2 for 10 minutes in the dark to remove endogenous peroxidase activity. After antigen retrieval by the citrate repair solution (pH = 6.0) in

a microwave oven for 10 min, the sections were sealed with goat non-immune serum (MXB Biotechnology Company, Fujian, China) and incubated with a primary antibody for MCM2 or SAE1 at a 1:400 or 1:800 dilution (Abcam, UK) overnight at 4°C. Following incubation with the secondary antibody, DAB kit (MXB Biotechnology Company, Fujian, China) was applied for the chromogenic reaction. The sections were then counterstained with hematoxylin (Beijing solarbio science & technology Company, China). The staining results were analyzed and scored independently by two experienced pathologists. Based on the staining intensity and the positive percentage of tumor cells, samples were scored as four grades from 0 to 3 (0, negative; 1, weakly positive; 2, moderately positive; 3, strong positive). A score of 2/3 was defined as high protein expression and a score of 0/1 was defined as low protein expression.

RESULTS

Global Profiling of LUSC Proteomics and Phosphoproteomics

To systematically investigate the characteristics of LUSC in China, LUSC tissues of 25 patients with tumor purity >50% were used to build proteomic and phosphoproteomic data (**Figure 1A**). After database searching by PD-SEQUEST HT (25, 32), 10,003 proteins were identified with high confidence (FDR<1%), in which 9,907 were quantified. On average, 8,516 proteins per sample were identified, 8,360 proteins per sample were quantified (**Figure 1B**). A total of 6,523 proteins were quantified in all samples (**Supplementary Table S2**). To control the variation among batches, the proteomic data of five internal references from five batches were assessed using Pearson correlation. The results displayed high correlations among batches with an average Pearson correlation coefficient of 0.98 (**Figure S1A**, **Supplementary Table S2**). In order to rule out samples with abnormal distribution of protein abundance, the density plot of $\log_2$-transformed TMT ratios of proteins and dip statistic (33) were used to characterize these distributions, which showed that 25 samples were unimodal (Gaussian or normal) distribution (**Figure S1B**, **Supplementary Table S3**). After data normalization, each sample had a similar distribution with $\log_2$-transformed TMT ratios centered at zero without batch effect (**Figure S1C**, **Supplementary Table S3**). By subcellular distribution analysis through Gene ontology database from PANTHER14.1 (34), we found the most identified proteins were nuclear proteins, followed by cytoplasmic proteins and organelle proteins, which was consistent with Reference Proteomic dataset and could reflect the real distribution in tissue without subcellular preference. Most interestingly, a large number of proteins (1,841 proteins) were annotated as cellular components in extracellular space, suggesting that proteome in LUSC could be related to the tumor microenvironment (**Figure 1C**, **Supplementary Table S4**).

For phosphoproteomic analysis with database searching by PD-SEQUEST HT (27), 13,721 unique phosphopeptides (FDR <1% and delta Cn < 0.05) and 9,601 unique phosphosites (localization probability > 0.75) on 3,786 proteins (FDR<1%) were identified (**Figure 1D**, **Supplementary Tables S5, S6**). The distribution of phosphoserine (pS), phosphothreonine (pT) and phosphotyrosine (pY) sites was 92.34% (8,866), 7.32% (703), and 0.33% (32), respectively (**Figure 1E**, **Supplementary Table S6**). Similarly, correlation of phosphosites in five internal references were assessed to control variation among batches, which displayed high correlation with average Pearson correlation coefficient 0.789 (**Figure S1D**, **Supplementary Table S7**). After data normalization, each sample had a similar phosphosite distribution with $\log_2$-transformed TMT ratios centered at zero without batch effect (**Figure S1E**, **Supplementary Table S7**).

Clustering Based on Protein Abundance

In the follow-up analysis, a sample with incomplete clinical information was removed, the remaining 24 samples were clustered by consensus clustering using k-means manner to explore the proteomic difference among LUSC tissues. Visually, the consensus matrix for $k = 2$ appeared to have the cleanest separation between clusters (**Figure 2A**, **Figure S2A**). The consensus CDF and delta area plot showed that there was no significant increase in the area under the consensus CDF as k increased from two (**Figure S2B**). Furthermore, the average silhouette distance for $k = 2$ (0.14) was larger than $k = 3$ (0.09). In addition, phosphoproteomic clustering analysis was also performed, in which the top 1,000 most varied phosphoproteins with less than 20% missing values within twenty-four tumors were used because the number of identified phosphoproteins were much less than the number of identified of proteins. The phosphoproteomic clustering was almost consistent with proteomic clustering (**Figure 2A**, **Figures S3A,B**). Therefore, 24 samples can be clustered to Cluster I ($n = 12$) and Cluster II ($n = 12$). The clustering results were then verified by principal component analysis, in which component 1 and component 2 accounted for 26.3% of the total data variation and the two components can distinguish Cluster I and Cluster II (**Figure 2B**).

For 6,523 proteins which were quantified in all the 24 samples (**Supplementary Table S8**), 435 proteins were significantly upregulated in Cluster I and 266 proteins were significantly upregulated in Cluster II (fold change> 1.5 and t-test p-value< 0.05) (**Figure 2C**). Gene set enrichment analysis (GSEA) showed that spliceosome pathway was enriched in Cluster I, focal adhesion and immune-related pathways (e.g. Complement and coagulation cascades) were enriched in Cluster II (**Figure 2D**, **Supplementary Table S9**). For spliceosome, various studies have highlighted the significance of altered RNA splicing in cancer (35). The leading edge proteins in the spliceosome pathway were major spliceosome components and expressed higher in Cluster I, including SF3B1 (splicing factor 3b subunit 1) which was identified with recurrent mutations in various cancer (**Figure 2E**) (36, 37). In focal adhesion, some of the leading-edge proteins were constituents which participate in the structure linking membrane receptors and the actin cytoskeleton, while others contribute to signal transduction, including protein kinases and phosphatases (e.g., PTEN) (**Figure 2F**). Immune-related pathways (e.g., complement and coagulation cascades) were also enriched in Cluster II, which was

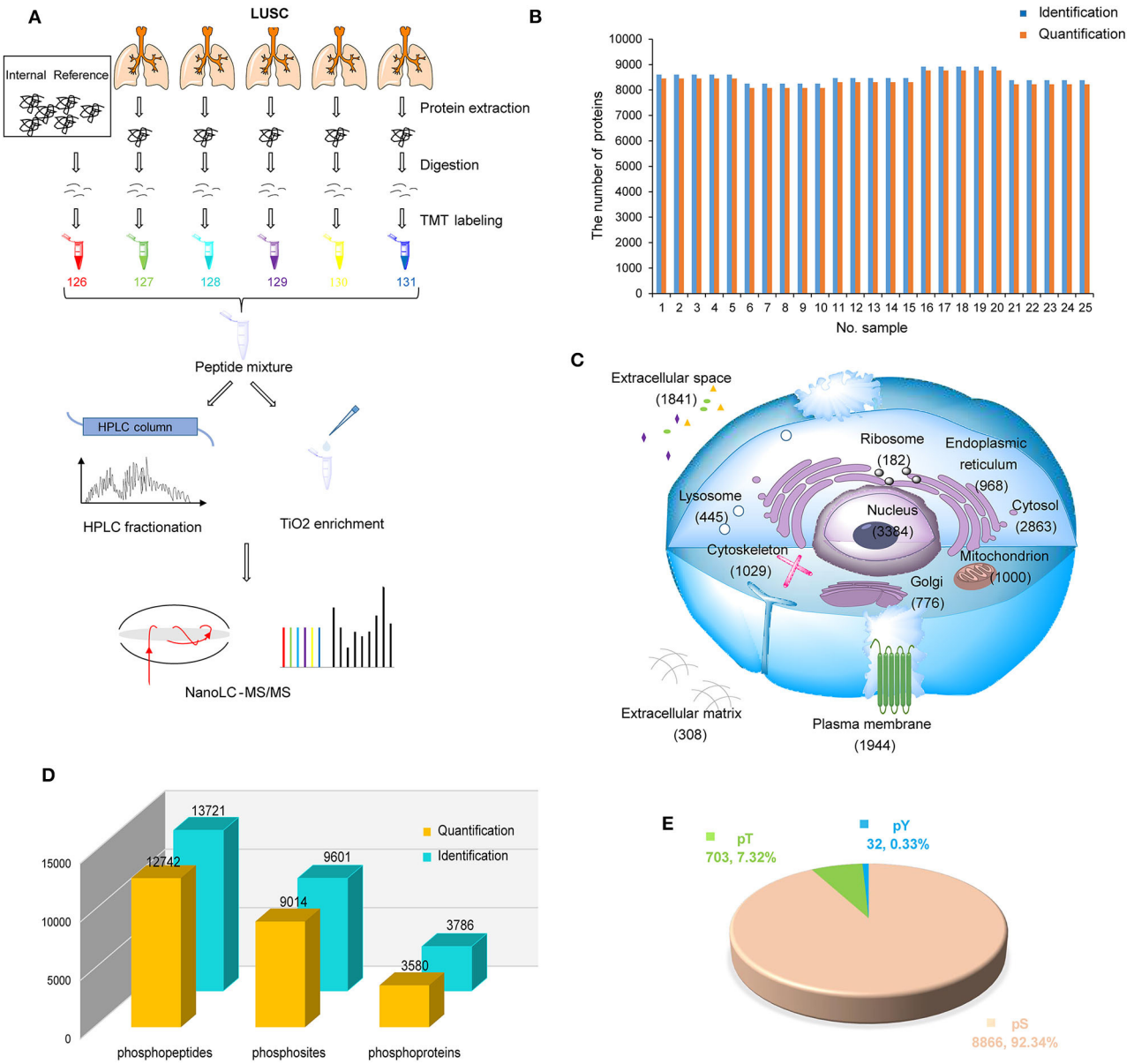


FIGURE 1 | Proteomic and phosphoproteomic analysis of lung squamous cell carcinoma (LUSC). **(A)** Workflow of proteomic and phosphoproteomic profiling of LUSC. **(B)** The number of identified and quantified proteins in proteomic profiling. **(C)** Subcellular distribution of LUSC gene products annotated with Gene Ontology. **(D)** The number of identified and quantified phosphopeptides/sites/proteins in phosphoproteomic profiling. **(E)** Phosphosite distribution on S/T/Y phosphorylation residues.

similar to "inflammatory" subtype in previous data reported in samples from Western countries (11).

Phosphoproteomic Analysis Based on Clustering

In order to further investigate the differences in signaling pathways between the two clusters, phosphoproteomic data including 7,973 phosphosites (at least two values in each cluster, **Supplementary Table S10**) was analyzed. The scatter plot showed that 329 phosphosites on 224 proteins were significantly upregulated in Cluster I and 333 phosphosites on 264 proteins were significantly upregulated in Cluster II (fold change> 1.5 and *t*-test *p*-value < 0.05) (**Figure 3A**). Then proteins with up/down-regulated phosphosites were used for KEGG pathway analysis via DAVID. Consistent with proteomic analysis, spliceosome was enriched in Cluster I, immune-related pathways and focal adhesion were enriched in Cluster II (**Figure 3B**, **Supplementary Table S11**). Besides, several proteins in Ras signaling pathway showed highly phosphorylated sites in Cluster II, including BAD Ser-134 (**Figure 3C**). BAD Ser-134 was reported to be phosphorylated by RAF, which leads to increased proliferation of cancer cells

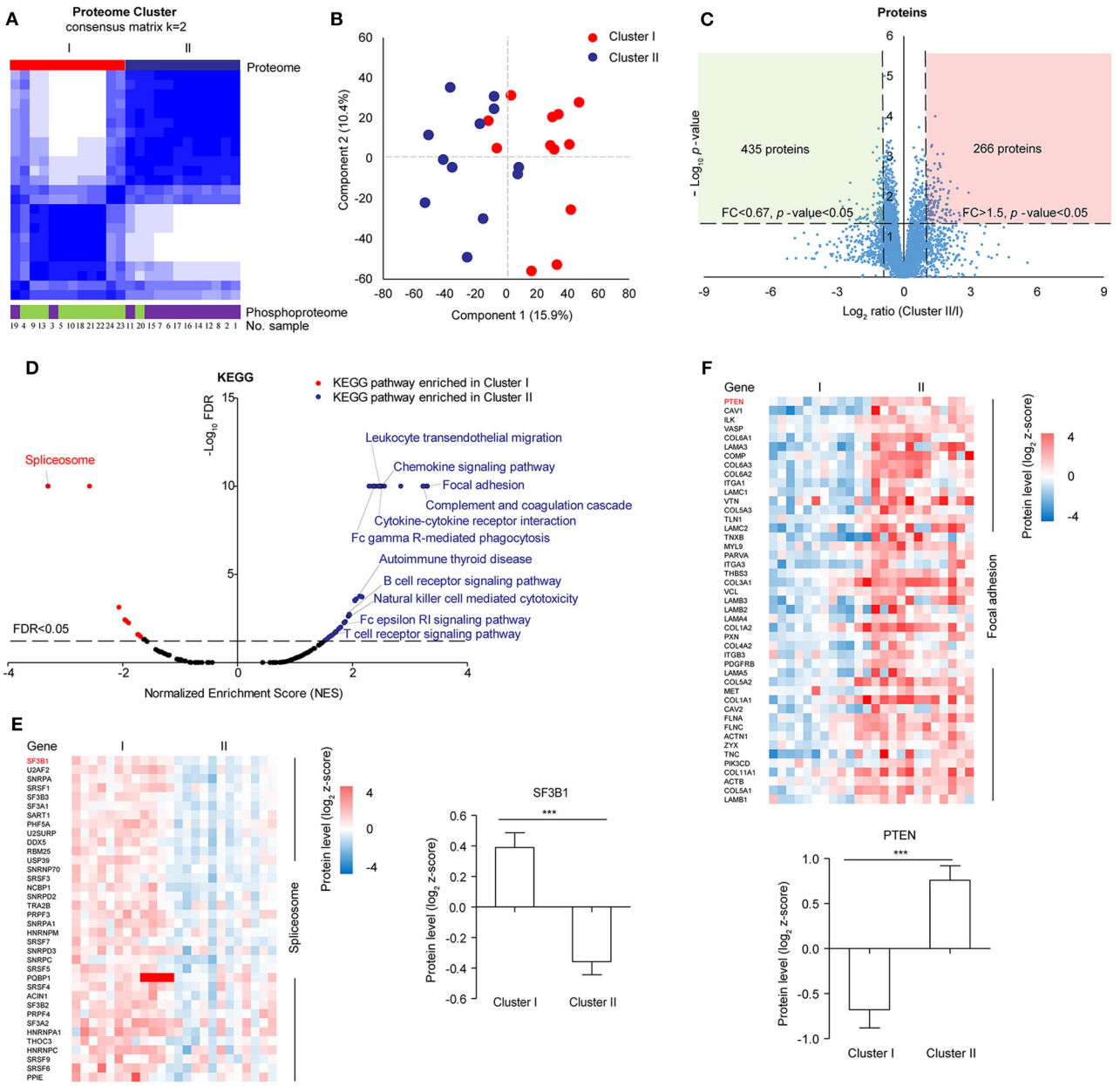


FIGURE 2 | Proteomic clustering of LUSC. **(A)** Consensus-clustering analysis of proteomic profiling using the top 1,000 most varied proteins with no missing values within twenty-four tumors. Consensus-clustering analysis of phosphomic profiling using the top 1,000 most varied phosphoproteins with missing values <20% within twenty-four tumors. **(B)** Principal Component Analysis (PCA) of two proteomic clusters. Red represented Cluster I, and blue represented Cluster II. **(C)** Scatter plot depicting the fold change of protein abundance comparing cluster II with cluster I. Log_2 fold changes were shown on the x-axis and $–log_{10}$ *p*-values were shown on the y-axis. The vertical dashed lines indicated fold change > 1.5 and the horizontal dashed line indicated *p*-value < 0.05 (*t*-test). **(D)** GSEA analysis of proteomic data between cluster I and cluster II. The scatter plot showed the enriched KEGG pathways from the Molecular Signatures Database (MSigDB). Normalized enrichment score (NES) was shown on the x-axis and $-log_{10}$ FDR was shown on the y-axis. The horizontal dashed line indicated FDR < 0.05. The labeled pathways were the most significant pathways or pathways consistent with previous data reported in samples from Western countries. **(E)** Differential protein expression of cluster I and cluster II in spliceosome. **(F)** Differential protein expression of cluster I and cluster II in focal adhesion.

(38). In the cascade of signaling pathway, kinase activity plays an important part by regulating the change of phosphosite level. Here, we found the potential change of activity in several kinases by Kinase-Substrate Enrichment Analysis (KSEA) (**Figure 3D**, **Supplementary Table S12**) (39, 40). Among these, MAPKAPK2/3/5 and CDK1/2 showed higher kinase activity in Cluster I, whereas EGFR, PRKCA/G, MAPK1/3/10/11/12/13, and MAPK2K1/4/7 showed higher kinase activity in Cluster II. Different kinases in MAPK pathway may be involved in different biological processes, for example, MAPKAPK family corresponds

to TNF stimulation and most of MAPK family corresponds to GF stimulation (41, 42). Further analysis discovered that several kinases were important components in MAPK signaling pathway and displayed in the scheme (**Figure 3E**).

Analysis of Lymph Node Metastasis (LNM)

Clinicopathologic characteristics are important factors which could be related to molecular features and patient prognosis. Therefore, Clinical information of the proteomic set samples was linked with clustering (**Supplementary Table S13**). Compared with Cluster II, Cluster I showed a little poor prognosis with log-rank p-value 0.2 within 3-year follow-up (**Figure S4A**). Although most clinicopathological characteristics showed no significant relationship with clustering statically, it was prominent that lymph node metastasis (LNM) showed some correlation with clustering (Fisher's exact test p-value 0.069) (**Figure 4A**). In detail, 6 of 7 patients without LNM (N0) were in cluster II and 11 of 17 patients with LNM (N1&2) were in cluster I. Patients with LNM showed a little poor prognosis (log-rank p-value 0.0645) within 3-years follow-up (**Figure 4B**, **Figure S5A**). Tumors often engage the lymphatic system to invade and metastasize (43). LNM is an indication of poor prognosis (44). Together, all the evidence indicated that the study on LNM in LUSC was valuable. To identify metastatic proteins, we took advantage of the Human Cancer Metastasis Database (45). There are 355 proteins associated with lung cancer metastasis in HCMDB, among which, 16 showed differential expression between N0 and N1&2 (fold change$>$ 1.5 and t-test p-value $<$ 0.05) (**Figure 4C**). In detail, 5 proteins [e.g., PRDX1 (peroxiredoxin 1) and ELAVL1 (ELAV like RNA binding protein 1)] were highly expressed in N1&2, whereas other 11 proteins (e.g., PTEN) were highly expressed in N0, indicating that these proteins may also play a role on LNM in LUSC.

In order to explore the pathways associated with LNM, the relative expression of 6,523 proteins (no missing value in 24 samples) were used to conduct KEGG analysis by GSEA. The result showed that immune-related pathways and focal adhesion were enriched in N0, cell cycle-related pathways and spliceosome were enriched in N1&2 (**Figure 4D**, **Supplementary Table S14**). To further investigate the differences in signaling pathways between N0 and N1&2, phosphoproteomic data including 7,209 phosphosites (at least two values in each cluster, **Supplementary Table S15**) was analyzed. The scatter plot showed that 260 phosphosites on 221 proteins were significantly upregulated in N0 and 186 phosphosites on 140 proteins were significantly upregulated in N1&2 (fold change $>$ 1.5 and t-test p-value $<$ 0.05) (**Figure 4E**). The proteins with differential phosphosites were used to enrich KEGG pathways via DAVID. Focal adhesion, as well as immune-related pathways, were enriched in N0, spliceosome was enriched in N1&2 (**Figure 4F**, **Supplementary Table S16**). We next explored upstream kinases enriched by the phosphosites in KSEA using the PhosphoSitePlus database (46) (**Supplementary Table S17**). This analysis revealed that 4 kinases (PRKACA, PRKCA, CSNK1A1 and BCR) showed higher kinase activity in LUSC without LNM, whereas the other 12 kinases showed higher kinase activity in LUSC with LNM, including cyclin-dependent kinase 1 (CDK1) and cyclin-dependent kinase 2 (CDK2) which are key regulatory enzymes in cell cycle, indicating that cell cycle may be potentially activated on the condition of LNM (**Figure 4G**).

Cell Cycle and DNA Replication in LUSC Based on LNM

Cancer is characterized by uncontrolled tumor cell proliferation resulting from aberrant activity of various cell cycle proteins. Therefore, cell cycle regulators are considered attractive targets in cancer therapy (47). In our data, cell cycle pathway was significantly enriched in the N1&2 group. The expressions of the leading edge proteins which contributed most to the enrichment score (ES) were upregulated with LNM (**Figure 5A**). Among these proteins, most were associated with DNA replication, as shown in the diagram (**Figure 5B**). DNA replication licensing factors- Minichromosome Maintenance (MCM) are essential for initiating and limiting DNA replication in cell cycle and implicates prognostic significance in lung cancer (48, 49). All members in the MCM protein family were highly expressed in the N1&2 group as well as in Cluster I (**Figure 5A**, **Figure S5C**). MCM2 is an independent predictor of survival in patients with non-small-cell lung cancer (50). To explore the influence of MCM2 on LUSC, we conducted IHC analysis based on TMA including 75 cases of LUSC patients, we found that MCM2 had higher expression in LUSC with LNM (**Figure 5C**, **Supplementary Table S18**). These findings suggested that high MCM2 expression in LUSC could be related to high biological malignant aggressiveness. Further survival analysis indicated that high expression of MCM2 was associated with poor prognosis (log-rank p-value $= 0.0489$) (**Figure 5D**). These results suggested that MCM2 might be a potential therapeutic target for LUSC, especially for LUSC with LNM.

DISCUSSION

LUSC accounts for a significant percentage of NSCLC (about 40%), but has limited biomarkers for diagnosis and therapy, except for a few immunotherapies targeting at PD1 and PD-L1 (6–8). In addition, systematic study on the molecular mechanism of Chinese LUSC patients is limited. Further study on the molecular mechanism of LUSC is needed for targeted therapy, especially in the Chinese cohort. To explore the protein expression pattern and the activation of signaling pathways in cancer, we investigated the proteomic and phosphoproteomic characteristics of LUSC from Chinese patients based on mass spectrometry analysis. In order to distinguish molecular characteristics, the samples were clustered into two parts on the basis of proteomic profiling. Compared with previous data reported in samples from Western countries (5, 11), our data showed that immune-related pathways were enriched in Cluster II, which was similar to the "inflammatory" subtype at protein level and "secretory" subtype at mRNA level. Even though the "redox" was not a major pattern in Cluster I, PSAT1 (phosphoserine aminotransferase), a potential target in the "redox" subtype, was

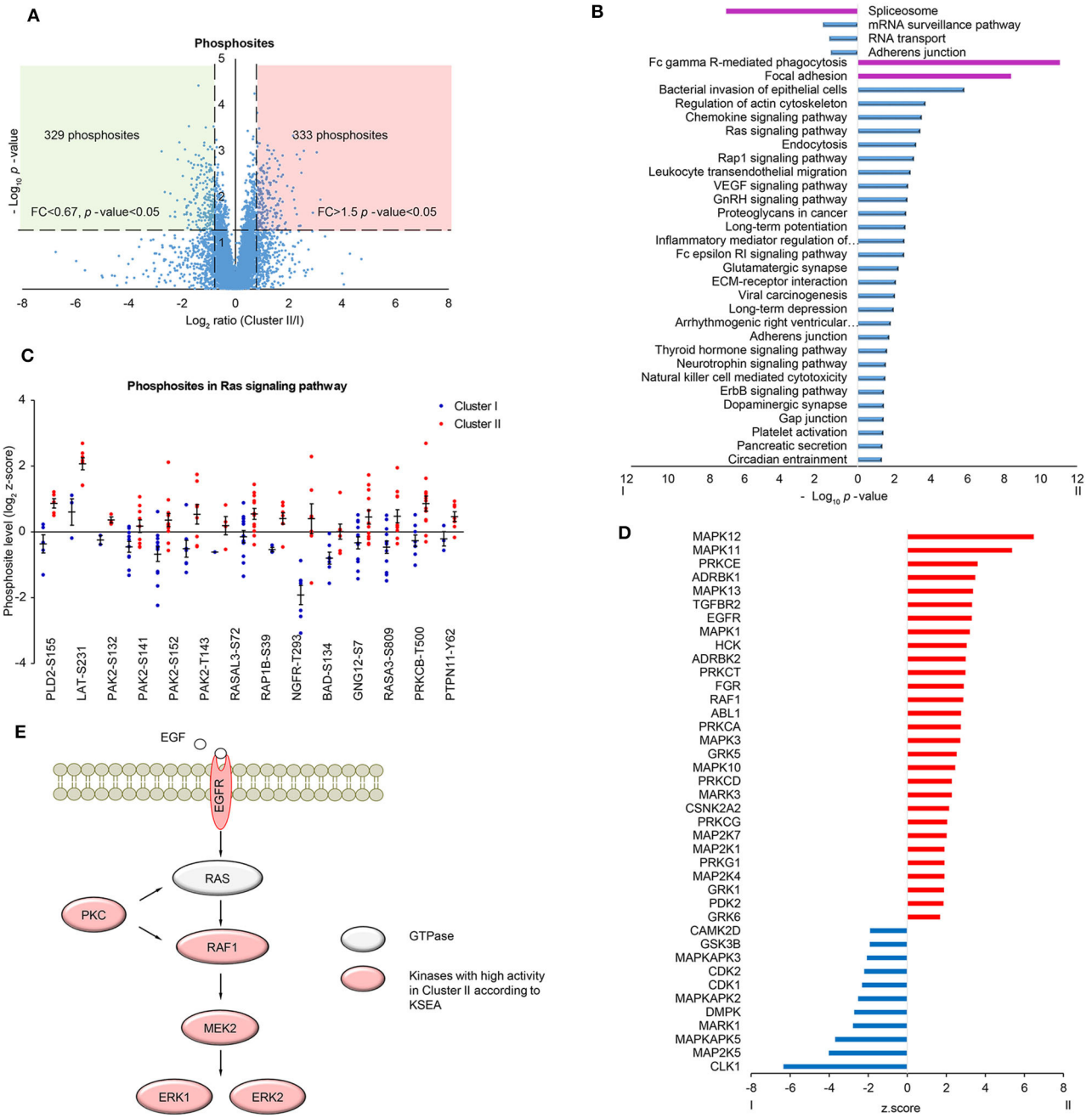


FIGURE 3 | Phosphoproteomic analysis of LUSC comparing cluster I with cluster II. **(A)** Scatter plot depicting the fold change in phosphosites comparing cluster II with cluster I. Log_2 fold changes were shown on the x-axis and $-log_{10}$ *p*-values are shown on the y-axis. The vertical dashed lines indicated fold change > 1.5 and the horizontal dashed line indicated *p*-value < 0.05 (*t*-test). **(B)** DAVID analysis of phosphoproteomic data between cluster I and cluster II (*p*-value < 0.05). The bar chart showed the enriched KEGG pathways of the differential phosphoproteins between cluster I and II (fold change > 1.5, *p*-value < 0.05). The highlighted pathways were consistent with those enriched in protein level (purple). **(C)** The quantification of phosphosites in Ras signaling pathway. Data were presented as dot plot with mean ± SEM. **(D)** Kinase-substrate enrichment analysis (KSEA) based on PhosphoSitePlus and NetworKIN database. Color blue/red for visual annotation of kinases that reached statistical significance (*p*-value < 0.05). **(E)** A scheme showing the enriched kinases in the MAPK pathway.

highly expressed in this cluster (**Figure S4B**) (11, 51). These results suggested that there were some similar patterns across the races. Beyond the previous study, our data indicated that spliceosome pathway was enriched in cluster I, focal adhesion pathway was enriched in cluster II in both proteomic and phosphoproteomic data.

In the pathway analysis of proteomic data, we found one-ubiquitin mediated proteolysis was enriched in Cluster I with

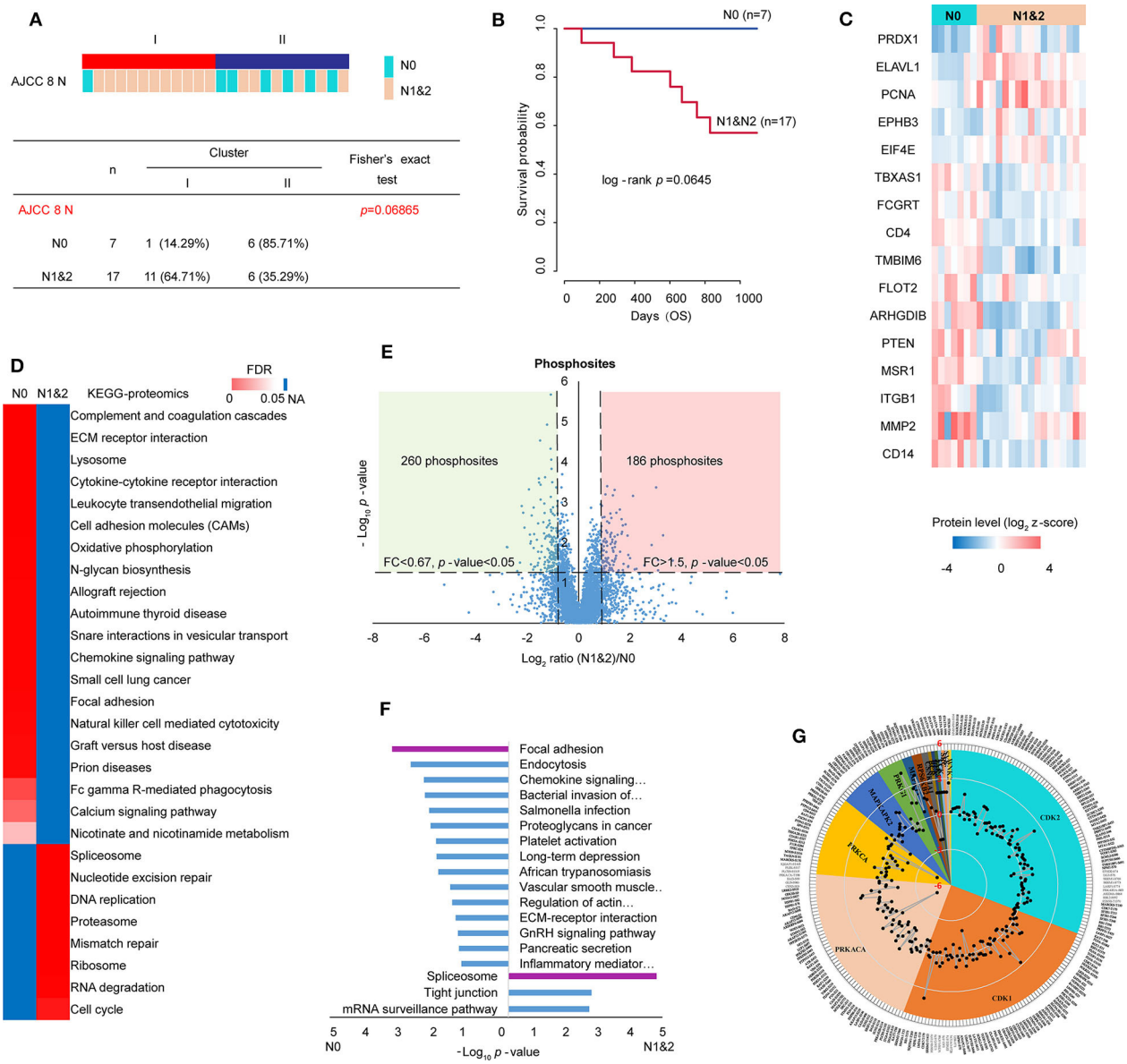


FIGURE 4 | Analysis of lymph node metastasis (LNM). **(A)** Relationship of cluster with LNM. Fisher's exact test was adopted for the analysis. **(B)** The prognosis of patients with or without LNM within 3-year follow-up. **(C)** 16 differential metastasis-related proteins based on total 355 proteins which were associated with lung cancer metastasis from HCMDB database (fold change> 1.5, p-value < 0.05). **(D)** GSEA analysis of proteomic data with and without LNM (tumor samples, $n = 24$): N0 (without LNM, $n = 7$) and N1&2 (with LNM, $n = 17$). The heatmap showed the enriched KEGG pathways from MSigDB. The color in heatmap was according to FDR, and the darkest blue represents N/A. **(E)** Scatter plot depicting the fold change in phosphosites comparing N1&2 with N0. Log_2 fold changes were shown on the x-axis and $-log_{10}$ p-values are shown on the y-axis. The vertical dashed lines indicated fold change > 1.5 and the horizontal dashed line indicated p-value < 0.05 (t-test). **(F)** DAVID analysis of phosphoproteomic data between N1&2 and N0. The bar chart showed the enriched KEGG pathways of the regulated phosphoproteins between N1&2 and N0 (fold change> 1.5, p-value < 0.05). The highlighted pathways were consistent with those enriched in the protein level. **(G)** Circular plot represented potential kinases and phosphosites retrieved from the PhosphoSitePlus database. The phosphosites were differentially expressed between N1&2 and N0 with a cutoff of 1.5-fold change. Different colors correspond to various kinases that were predicted as upstream regulators of the phosphosites (p-value < 0.05). The outer circle shows the sites. The radar map shows the fold changes of the sites.

FDR 0.18 (**Figure S4C**), and PPI network of the leading edge proteins analyzed by STRING revealed that most proteins were closely associated with the surrounding proteins in the network. Among these proteins, SAE1 was reported to be highly expressed in a variety of cancers and promotes tumor progression as well as poor prognosis (52–55). To explore the influence of SAE1 on LUSC, we conducted IHC analysis based on TMA including 75 cases of LUSC patients, which showed that high expression of SAE1 was associated with poor prognosis (**Figure S4D**, **Supplementary Table S16**). In the

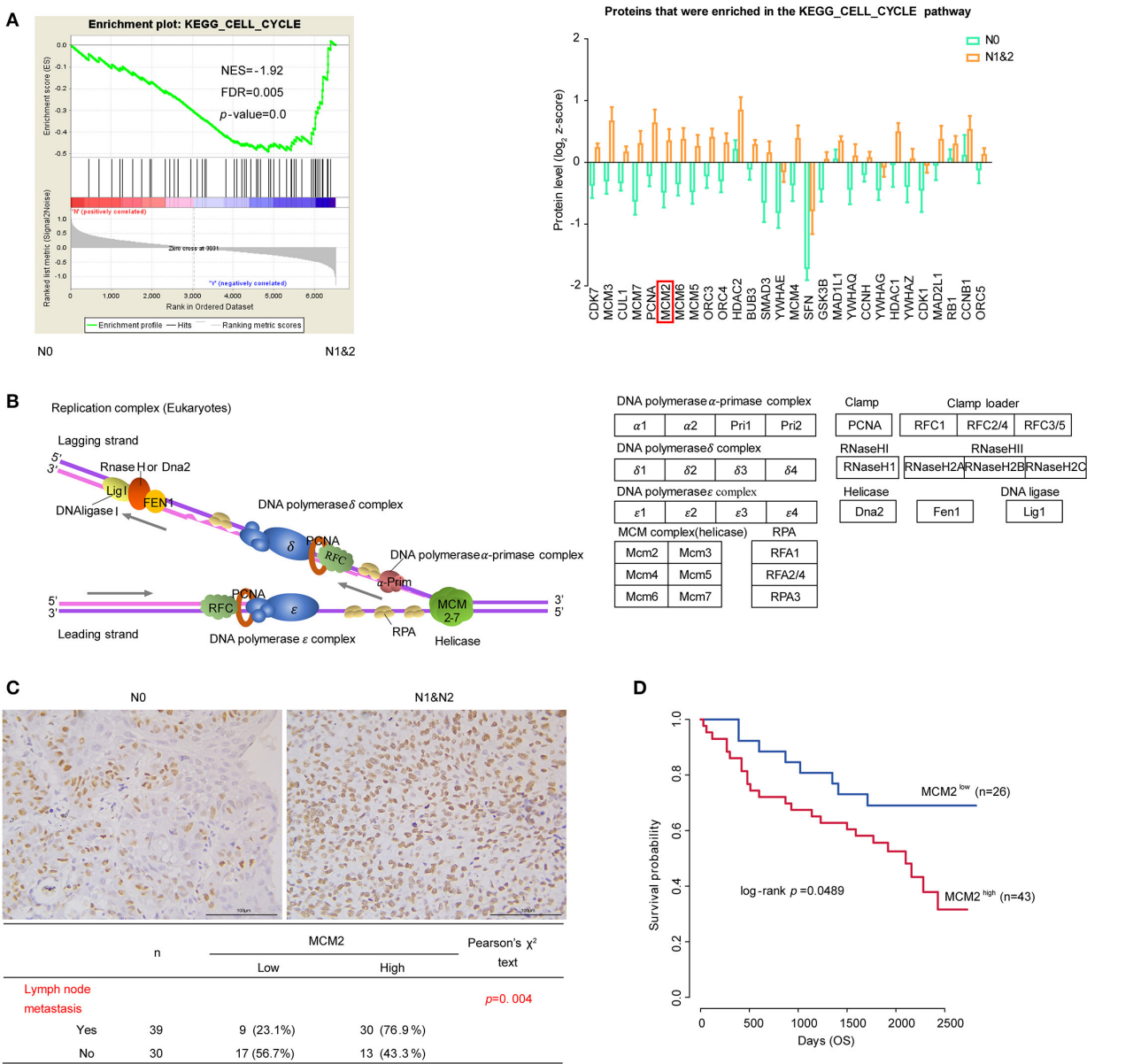


FIGURE 5 | Cell cycle and DNA replication in LUSC based on LNM. **(A)** Cell cycle pathway and proteins that were enriched in the pathway. The protein quantifications are presented as mean ± SEM. **(B)** The diagram of DNA replication in cell cycle. Different protein complexes are represented in different colors. **(C)** The relationship MCM2 protein expression and LNM in tissue microarray (TMA) assay. **(D)** MCM2 expression and survival analysis in tissue microarray (TMA) assay. Chi-square test was adopted for the correlation analysis, and Kaplan–Meier plot (Log-rank test) was adopted to describe overall survival (OS) analysis.

pathway analysis of phosphoproteomic data, Ras signaling pathway was highly activated in Cluster II and the activity of several kinases was shown to be elevated in Cluster II compared with Cluster I, indicating Ras signaling pathway may be different in these two clusters.

In order to comprehensively analyze molecular features and clinicopathologic characteristics, we first linked clinicopathologic characteristics with clustering, which showed some correlation between LNM and clustering. Patients with LNM showed a little poor prognosis. Considering the lymphatic system was often engaged by tumor invasion and metastasis (43), we supposed that LNM should have a significant influence on prognosis with a larger sample size, which was confirmed by analyzing public data from the previous study (**Figure S5B**) (11). Both proteomic and phosphoproteomic data were used for pathway analysis and showed enrichment of several major pathways in N0 (or N1&2), similar to those found in Cluster II (or Cluster I). Besides, key regulatory enzymes in cell cycle were also enriched in LUSC with LNM by KSEA.

DNA replication is a key in cell cycle and MCM2 was discovered with high expression in patients with LNM by proteomics analysis and in patients with poor prognosis by IHC analysis. MCM2 is one of the members of the MCM protein family. It forms MCM complex with its family members MCM3-7. The MCM complex is a replication helicase, which is essential for the DNA replication initiation and extension of cell cycle in eukaryotic cells (56). The expression level of MCM proteins (MCMs) in normal cells changes with the progress of the cell cycle. In the G1 phase of the cell cycle, CDK activates the transcription factor E2F by phosphorylating RB, E2F can combine with the promoter region of MCMs to promote its transcription (57, 58). In senescent cells, p53 can synthesize microRNAs to degrade MCMs mRNA (59). In summary, MCM proteins could be dysregulated by different signaling pathways in cancer. In our proteomic data, cyclin-dependent kinases (CDK1/2) showed higher kinase activity in LUSC with LNM which may regulate MCM2 expression. However, The Cancer Genome Atlas (TCGA) dataset showed the RNA level of MCM2 was not correlated with prognosis in LUSC (**Figure S5D**) (6), which suggested there may be different features in genomics and proteomics. Further, MCM2 was shown to be not correlated with prognosis in LUSC from Western patients (**Figure S5E**) (11), which indicated there may exist different mechanisms among races. In our IHC analysis, MCM2 displayed higher expression in advanced patients (**Figure S5F**). Altogether, our study provided a proteomic and phosphoproteomic data resource about LUSC from Chinese patients, which could give several clues on potentially targeted proteins for precision medicine.

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: the integrated Proteome resources (iProx database) (http://www.iprox.org/index), accession number IPX0001833000.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Cancer Institute/Hospital, Peking Union Medical College & Chinese Academy of Medical Sciences and Beijing Xuanwu Hospital. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

LP, XW, and LY conducted most of the proteomic and biochemical experiments with the help of YY, YM, and SC under the supervision of TX and MT. LP, XW, and LY analyzed the data with the help of LZhao, LZhai, and JX under the supervision of TX and MT. LP and XW wrote most of the manuscript with the guidance and help of TX and MT revised it. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Natural Science Foundation of China (Grant Nos. 81872888 and 81821005), the National Basic Research Program of China (973 Program) (Grant No. 2014CBA02004), the National Science and Technology Major Project Key New Drug Creation and Manufacturing Program (Grant No. 2018ZX09711002-004), the CAMS Innovation Fund for Medical Sciences (CIFMS) of China (Grant Nos. 2016-I2M-1-001 and 2019-I2M-1-003), the National Key Research and Development Program of China (Grant No. 2017YFC0908401), the Innovation research team of high-level local universities in Shanghai (Grant No. SSMU-ZDCX20181202), and K. C. Wong Education Foundation.

ACKNOWLEDGMENTS

We would like to thank all the participants of this work. We also thank the Beijing Xuanwu Hospital for its assistance in specimen collection. We also acknowledge NCCN for the Clinical Practice Guidelines in Oncology (www.nccn.org).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.00963/full#supplementary-material

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* (2013) 63:11–30. doi: 10.3322/caac.21166
2. Fitzmaurice C, Abate D, Abbasi N, Abbastabar H, Abd-Allah F, Abdel-Rahman O, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-Adjusted life-Years for 29 cancer groups, 1990 to 2017: a Systematic analysis for the global burden of disease study. *JAMA Oncol.* (2019) 5:1749–68. doi: 10.1001/jamaoncol.2019.2996
3. Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer.* (2014) 14:535–46. doi: 10.1038/nrc3775
4. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet.* (2016) 48:607–16. doi: 10.1038/ng.3564
5. Wilkerson MD, Yin X, Hoadley KA, Liu Y, Hayward MC, Cabanski CR, et al. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin Cancer Res.* (2010) 16:4864–75. doi: 10.1158/1078-0432.CCR-10-0199
6. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers. *Nature.* (2012) 489:519–25. doi: 10.1038/nature11404
7. Network CGAR. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* (2014) 511:543–50. doi: 10.1038/nature13385

8. National Comprehensive Cancer Network. *(NCCN) Clinical Practice Guidelines in Oncology. Non-Small Cell Lung Cancer* (2019). Available online at: www.nccn.org
9. Olena M, Nicole M, White A, DeSouza LV, Krakovska O, Matta A, et al. Quantitative proteomic analysis in metastatic renal cell carcinoma reveals a unique set of proteins with potential prognostic significance. *Mol Cell Prot.* (2013) 12:132–44. doi: 10.1074/mcp.M112.020701
10. Houdt WV, Emmink B, Pham T, Piersma SR, Verheem A, Vries RG, et al. Comparative proteomics of colon cancer stem cells and differentiated tumor cells identifies bIRC6 as a potential therapeutic target. *Mol Cell Prot.* (2011) 10:1–11. doi: 10.1074/mcp.M111.011353
11. Stewart PA, Welsh EA, Slebos RJC, Fang B, Izumi V, Chambers M, et al. Proteogenomic landscape of squamous cell lung cancer. *Nat Commun.* (2019) 10:3578. doi: 10.1038/s41467-019-11452-x
12. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* (2020) 70:7–30. doi: 10.3322/caac.21590
13. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* (2015) 65:87–108. doi: 10.3322/caac.21262
14. Chen W, Sun K, Zheng R, Zeng H, Zhang S, Xia C, et al. Cancer incidence and mortality in china, 2014. *Chin J Cancer Res.* (2018) 30:1–12. doi: 10.21147/j.issn.1000-9604.2018.01.01
15. Wang C, Yin R, Dai J, Gu Y, Cui S, Ma H, et al. Whole-genome sequencing reveals genomic signatures associated with the inflammatory microenvironments in chinese nSCLC patients. *Nat Communicat.* (2018) 9:2054. doi: 10.1158/1538-7445.AM2018-5375
16. Zeng GQ, Zhang PF, Deng X, Yu FL, Li C, Xu Y, et al. Identification of candidate biomarkers for early detection of human lung squamous cell cancer by quantitative proteomics. *Mol Cell Prot.* (2012) 11:M111 013946. doi: 10.1074/mcp.M111.013946
17. Clark DJ, Dhanasekaran SM, Petralia F, Pan J, Song X, Hu Y, et al. Integrated proteogenomic characterization of clear cell renal cell carcinoma. *Cell.* (2019) 179:964–83. doi: 10.1158/1538-7445.AM2019-2458
18. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell.* (2019) 177:1035–49. doi: 10.1016/j.cell.2019.03.030
19. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, et al. Integrated proteogenomic characterization of hBV-Related hepatocellular carcinoma. *Cell.* (2019) 179:561–77. doi: 10.1016/j.cell.2019.08.052
20. Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. *Nature.* (2019) 567:257–61. doi: 10.1038/s41586-019-0987-8
21. Coscia F, Lengyel E, Duraiswamy J, Ashcroft B, Bassani-Sternberg M, Wierer M, et al. Multi-level proteomics identifies cT45 as a chemosensitivity mediator and immunotherapy target in ovarian cancer. *Cell.* (2018) 175:159–70. doi: 10.1016/j.cell.2018.08.065
22. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature.* (2016) 534:55–62. doi: 10.1038/nature18003
23. Drake JM, Paull EO, Graham NA, Lee JK, Smith BA, Titz B, et al. Phosphoproteome integration reveals patient-Specific networks in prostate cancer. *Cell.* (2016) 166:1041–54. doi: 10.1016/j.cell.2016.07.007
24. Kettenbach AN, Gerber SA. Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. *Anal Chem.* (2011) 83:7635–44. doi: 10.1021/ac201894j
25. Adav SS, Wei J, Terence Y, Ang BCH, Yip LWL, Sze SK. Proteomic analysis of aqueous humor from primary open angle glaucoma patients on drug treatment revealed altered complement activation cascade. *J Prot Res.* (2018) 17:2499–510. doi: 10.1021/acs.jproteome.8b00244
26. Marina Spivak JW, Le on Bottou LK, William Stafford Noble. Improvements to the percolator algorithm for peptide identification from shotgun proteomics data sets. *J Prot Res.* (2009) 8:3737–45. doi: 10.1021/pr801109k
27. Wu XY, Zahari MS, Renuse S, Nirujogi RS, Min-Sik K, Srikanth SM, et al. Phosphoproteomic Analysis Identifies Focal Adhesion Kinase 2 (FAK2) as a Potential Therapeutic Target for Tamoxifen Resistance in Breast Cancer. *Mol Cell Prot.* (2015) 14:2887–900. doi: 10.1074/mcp.M115.050484
28. Monti S, Tamayo P, Mesirov J, GOLUB T. Consensus clustering_ a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn.* (2003) 52:91–118.
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
30. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* (2007) 35(Web Server issue):169–75. doi: 10.1093/nar/gkm415
31. Ogata H, Goto S, Sato K, Wataru Fujibuchi HB, Kanehisa* aM. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (1999) 27:29–34. doi: 10.1093/nar/27.1.29
32. Plubell DL, Wilmarth PA, Zhao Y, Fenton AM, Minnier J, Reddy AP, et al. Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Mol Cell Prot.* (2017) 16:873–90. doi: 10.1074/mcp.M116.065524
33. Hartigan J, Hartigan P. The dip test of unimodality. *Ann Stat.* (1985) 13:70–84. doi: 10.1214/aos/1176346577
34. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new pANTHER gO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* (2019) 47:D419–D26. doi: 10.1093/nar/gky1038
35. Joana D, Pedro BG, Maria CF. Targeting mRNA processing as an anticancer strategy. *Nat Rev Drug Dis.* (2019) 19:112–29. doi: 10.1038/s41573-019-0042-3
36. Wang L, Lawrence MS, Wan Y, Petar Stojanov BA, Carrie Sougnez BS, Kristen Stevenson MS, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med.* (2011) 365:2497–506. doi: 10.1056/NEJMoa1109016
37. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor sF3B1 gene in chronic lymphocytic leukemia. *Nature genetics.* (2011) 44:47–52. doi: 10.1038/ng.1032
38. Polzien L, Baljuls A, Albrecht M, Hekman M, Rapp UR. BAD contributes to rAF-mediated proliferation and cooperates with b-RAF-V600E in cancer signaling. *J Biol Chem.* (2011) 286:17934–44. doi: 10.1074/jbc.M110.177345
39. Wiredja DD, Koyuturk M, Chance MR. The kSEA app: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics.* (2017) 33:3489–91. doi: 10.1093/bioinformatics/btx415
40. Casado P, Rodriguez-Prados JC, Cosulich SC, Sylvie Guichard, Bart Vanhaesebroeck SJ, Pedro R, et al. Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci Signal.* (2013) 6:rs6. doi: 10.1126/scisignal.2003573
41. Ronkina N, Menon MB, Schwermann J, Tiedje C, Hitti E, Kotlyarov A, et al. MAPKAP kinases mK2 and mK3 in inflammation: complex regulation of tNF biosynthesis via expression and phosphorylation of tristetraprolin. *Biochem Pharm.* (2010) 80:1915–20. doi: 10.1016/j.bcp.2010.06.021
42. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell.* (2010) 141:1117–34. doi: 10.1016/j.cell.2010.06.011
43. Swartz MA, Lund AW. Lymphatic and interstitial flow in the tumour microenvironment: linking mechanobiology with immunity. *Nat Rev Cancer.* (2012) 12:210–9. doi: 10.1038/nrc3186
44. De Boer M, Van Dijck JA, Bult P, Borm GF, Tjan-Heijnen VC. Breast cancer prognosis and occult lymph node metastases, isolated tumor cells, and micrometastases. *J Natl Cancer Inst.* (2010) 102:410–25. doi: 10.1093/jnci/djq008
45. Zheng G, Ma Y, Zou Y, Yin A, Li W, Dong D. HCMDB: the human cancer metastasis database. *Nucleic Acids Res.* (2018) 46:950–5. doi: 10.1093/nar/gkx1008
46. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, pTMs and recalibrations. *Nucleic Acids Res.* (2015) 43(Database issue):512–20. doi: 10.1093/nar/gku1267
47. Otto T, Sicinski P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer.* (2017) 17:93–115. doi: 10.1038/nrc.2016.138
48. Tye BK. MCM proteins in dNA replication. *Annu Rev Biochem.* (1999) 68:649–86. doi: 10.1146/annurev.biochem.68.1.649

49. Liu YZ, Wang BS, Jiang YY, Cao J, Hao JJ, Zhang Y, et al. MCMs expression in lung cancer: implication of prognostic significance. *J Cancer.* (2017) 8:3641–7. doi: 10.7150/jca.20777
50. Ramnath N, Hernandez FJ, Tan DF, Huberman JA, Natarajan N, Beck AF, et al. MCM2 is an independent predictor of survival in patients with non-small-cell lung cancer. *J Clin Oncol.* (2001) 19:4259–66. doi: 10.1200/JCO.2001.19.22.4259
51. DeNicola GM, Chen PH, Mullarky E, Sudderth JA, Hu Z, Wu D, et al. NRF2 regulates serine biosynthesis in non-small cell lung cancer. *Nat Gen.* (2015) 47:1475–81. doi: 10.1038/ng.3421
52. Kessler JD, Kahle KT, Sun T, Meerbrey KL, Schlabach MR, Schmitt EM, et al. A SUMOylation-dependent transcriptional subprogram is required for myc-driven tumorigenesis. *Science.* (2012) 335:348–53. doi: 10.1126/science.1212728
53. Tefano A, Miriam Lubrano Lavadera, Giacomo Di Palo, Majello B. SUMO-activating sAE1 transcription is positively regulated by myc. *Am J Cancer Res.* (2012) 2:330–4.
54. Yang Y, Liang Z, Xia Z, Wang X, Ma Y, Sheng Z, et al. SAE1 promotes human glioma progression through activating aKT sUMOylation-mediated signaling pathways. *Cell Commun Signal.* (2019) 17:82. doi: 10.1186/s12964-019-0392-9
55. Song N, Gu XD, Wang Y, Chen ZY, Shi LB. Lentivirus-mediated siRNA targeting sAE1 induces cell cycle arrest and apoptosis in colon cancer cell rKO. *Mol Biol.* (2014) 48:91–8. doi: 10.1134/S0026893314010129
56. Masai H, You Z, Arai K. Control of dNA replication: regulation and activation of eukaryotic replicative helicase, mCM. *IUBMB Life.* (2005) 57:323–35. doi: 10.1080/15216540500092419
57. Fei L, Xu H. Role of mCM2-7 protein phosphorylation in human cancer cells. *Cell Biosci.* (2018) 8:43. doi: 10.1186/s13578-018-0242-2
58. Ishimi Y. Regulation of mCM2-7 function. *Genes Genet Syst.* (2018) 93:125–33. doi: 10.1266/ggs.18-00026
59. Bai G, Smolka MB, Schimenti JC. Chronic dNA replication stress reduces replicative lifespan of cells by tRP53-Dependent, microRNA-Assisted mCM2-7 downregulation. *PLoS Genet.* (2016) 12:e1005787. doi: 10.1371/journal.pgen.1005787

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 © 2020 Pan, Wang, Yang, Zhao, Zhai, Xu, Yang, Mao, Cheng, Xiao and Tan. 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 Oncology

ORIGINAL RESEARCH
published: 03 July 2020
doi: 10.3389/fonc.2020.01072

Proteomic Analysis of the Air-Way Fluid in Lung Cancer. Detection of Periostin in Bronchoalveolar Lavage (BAL)

Yangying Zhou[1], *Weiming Yang*[1], *Minghui Ao*[1], *Naseruddin Höti*[1], *Edward Gabrielson*[1,2], *Daniel W. Chan*[1,2], *Hui Zhang*[1,2] *and Qing Kay Li*[1,2]*

[1] Department of Pathology, Johns Hopkins Medical Institutions, Baltimore, MD, United States, [2] Department of Oncology, Sidney Kimmel Cancer Center at Johns Hopkins Medical Institutions, Baltimore, MD, United States

OPEN ACCESS

Edited by:
Harsha Gowda,
The University of Queensland, Australia

Reviewed by:
Eswari Dodagatta-Marri,
University of California, San Francisco, United States
Raphael Carmo Valente,
Rio de Janeiro State University, Brazil

***Correspondence:**
Qing Kay Li
qli23@jhmi.edu

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 08 January 2020
Accepted: 28 May 2020
Published: 03 July 2020

Citation:
Zhou Y, Yang W, Ao M, Höti N, Gabrielson E, Chan DW, Zhang H and Li QK (2020) Proteomic Analysis of the Air-Way Fluid in Lung Cancer. Detection of Periostin in Bronchoalveolar Lavage (BAL). Front. Oncol. 10:1072. doi: 10.3389/fonc.2020.01072

Background: Bronchoalveolar lavage (BAL) is a specific type of air-way fluid. It is a commonly used clinical specimen for the diagnosis of benign diseases and cancers of the lung. Although previous studies have identified several disease-associated proteins in the BAL, the potential utility of BAL in lung cancer is still not well-studied. Based upon the fact that the majority of secreted proteins are glycoproteins, we have profiled *N*-glycoproteins in BAL collected from lung cancers, and investigated the expression of glycoproteins such as the matrix *N*-glycoprotein, periostin, in lung cancers.

Methods: BAL specimens ($n = 16$) were collected from lung cancer patients, and analyzed using mass spectrometry-based quantitative *N*-glycoproteomic technique. Additional BAL specimens ($n = 39$) were independently collected to further evaluate the expression of periostin by using an enzyme-linked immunosorbent assay (ELISA).

Results: A total of 462 glycoproteins were identified in BAL samples using *N*-glycoproteomic technique, including 290 in lung adenocarcinoma (ADC, $n = 5$), 376 in squamous cell carcinoma (SQCC, $n = 4$), 309 in small cell lung carcinoma (SCLC, $n = 4$), and 316 in benign lung disease ($n = 3$). The expressions of several glycoproteins were elevated, including 8 in ADC, 12 in SQCC, and 17 in SCLC, compared to benign BALs. The expression of periostin was detected in all subtypes of lung cancers. To further investigate the expression of periostin, an ELISA assay was performed using additional independently collected BALs ($n = 39$) The normalized levels of periostin in benign disease, ADC, SQCC, and SCLC were 255 ± 104 (mean ± SE) and 4,002 ± 2,181, 3,496 ± 1,765, and 1,772 ± 1,119 ng/mg of total BAL proteins.

Conclusion: Our findings demonstrate that proteomic analysis of BAL can be used for the study of cancer-associated extracellular proteins in air-way fluid from lung cancer patients.

Keywords: lung cancer, proteomic analysis, *N*-glycoprotein, bronchoalveolar lavage, periostin expression, ELISA

INTRODUCTION

Lung cancer is a heterogeneous group of tumors, including non-small cell carcinoma (NSCLC) and small cell carcinoma (SCLC). In the NSCLC, squamous cell carcinoma (SQCC) and adenocarcinoma (ADC) are the two dominant subtypes (1, 2). The recent large scale of genetic studies, including the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA), have identified multiple driver gene mutations in NSCLC (3, 4). These studies have greatly improved our molecular understanding of lung cancers and facilitated the development of targeted drugs and immunotherapies (3–9). However, despite these advancements, lung cancer remains one of the leading causes of cancer-associated mortality in the United States and worldwide (1, 2). Clinically, the majority of lung cancer patients present with the advanced-stage disease; thus, it is critical to improve detection and monitor disease progression in order to improve the overall survival of lung cancer.

Currently, we commonly use the combination of imaging and histological assessment of the lesion for the diagnosis of lung cancer. Several types of specimens can be used for histomorphological evaluation, such as bronchial brushing and bronchoalveolar lavage (BAL) samples, and transbronchial fine-needle aspiration (EBUS-TBNA) biopsy with or without ultrasound guidance specimens (10–13). In addition, biomolecular markers in the blood, bronchial epithelium, and sputum are also being tested (14, 15). Studies have shown that the use of glycoproteins could potentially improve the accuracy and specificity of lung cancer diagnosis (16, 17).

The air-way fluid bronchoalveolar lavage (BAL) contains proteins released from lung parenchyma that may be useful for the detection and monitory of certain benign lung diseases, as well as lung cancer progression (14–16, 18). In our previous study of lung ADC, we found that certain glycoproteins are differentially expressed between cancer and benign BAL using the solid-phase extraction of *N*-glycoprotein (SPEG) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) (16). The study showed that glycoproteins could be detected and quantified in BAL fluids, supporting that the glycoproteomic analysis of BAL could be useful for the identification of extracellular cancer-associated proteins, especially when the specimens did not qualify the criteria to have enough cells for histological examination (16–18). Therefore, analysis of *N*-glycoproteins in the air-way fluid of lung cancers has a potential utility for the identification of cancer-associated glycoproteins.

Studies have shown that tumor microenvironment, i.e., desmoplastic stroma, plays a critical role in tumor progression, invasion, and metastasis (19, 20). Several desmoplastic stroma *biomarker*s have been characterized in cancers (19, 20). Among these stromal biomarkers, periostin stands out in playing a remarkable role during cancer progression (21). Periostin is an extracellular matrix *N*-glycoprotein in the desmoplastic stroma around solid tumors, and is involved in the regulation of extracellular matrix (ECM) network (21). It interacts with other ECM proteins, such as collagen V, fibronectin, and tenascin-C to form an extracellular environment, in which cells can interact with each other to promote growth and survival. Periostin also interacts with various cell-surface receptors, particularly with integrins, through the PI3K/AKT signaling pathway to promote epithelial-mesenchymal transition (EMT), cell growth, and tumor angiogenesis and invasion (22, 23).

The expression of periostin in the desmoplastic stroma of the tumor has been associated with the aggressiveness of tumor behavior, increased metastatic potential, and advanced tumor stages in multiple cancer types, such as head and neck (24), colon (22), breast (25), lung (23, 26–30), and others (31). Previously, we have identified that the expression of periostin in the stroma of prostate cancer is associated with aggressive clinical behavior of the tumor (32). Furthermore, periostin is an extracellular protein and accessible via the bloodstream, as such, it has been considered to be a promising therapeutic target for cancer patients (31). In lung carcinoma, the expression of periostin has been linked to tumor invasiveness and metastasis (23, 26–30). Recent studies have also shown that periostin can be detected in the serum of NSCLC patients (29, 30), indicating the possible utility of periostin as a cancer-associated glycoprotein in lung cancer. However, the functional significances of *N*-glycoproteins expression in lung cancers, particularly periostin expression in air-way fluid, is still poorly understood.

In this study, we investigated the profile of *N*-glycoproteins in BAL specimens and identified periostin to be significantly overexpressed in different subtypes of lung cancers. These findings were further validated using an independent cohort of BAL specimens by enzyme-linked immunosorbent assay (ELISA).

MATERIALS AND METHODS

Collection of Lung BAL Specimens

The discarded BAL specimens were collected from the cytologic laboratory after the evaluation of lung cells. Briefly, 5–10 cc normal saline was applied to rinse the lung during the bronchoscopic procedure, the bronchoscopy was used to collect the fluids into a sterilized test tube and sent to the cytological laboratory. After carefully examined the pulmonary cells, the BAL fluids of ADC ($n = 5$), SQCC ($n = 4$), SCLC ($n = 4$), and benign ($n = 3$) were used for the proteomic analysis. For a validation study, additional BAL specimens were collected for the ELISA assay, including six cases of benign lung diseases, 18 cases of primary lung ADCs, 6 cases of primary lung SCLCs, and 9 cases of primary lung SQCCs. To remove the cellular debris, all BAL specimens were centrifuged and were then stored at −80°C freezer before the analysis.

This study was approved by the Institution Review Board (IRB) of Johns Hopkins Medical Center.

Peptide Extraction From BAL Specimens

All the BAL specimens were sonicated on an ice bath for 5 min and then centrifuged at 12,600 rpm for 10 min. One-hundred microliters of trifluoroethanol (TFE) was added into each specimen and incubated for 2 h. The protein concentration of the BAL samples was calculated using the BCA protein assay kit (Thermo Fisher Scientific Inc., Rockford, IL). Then

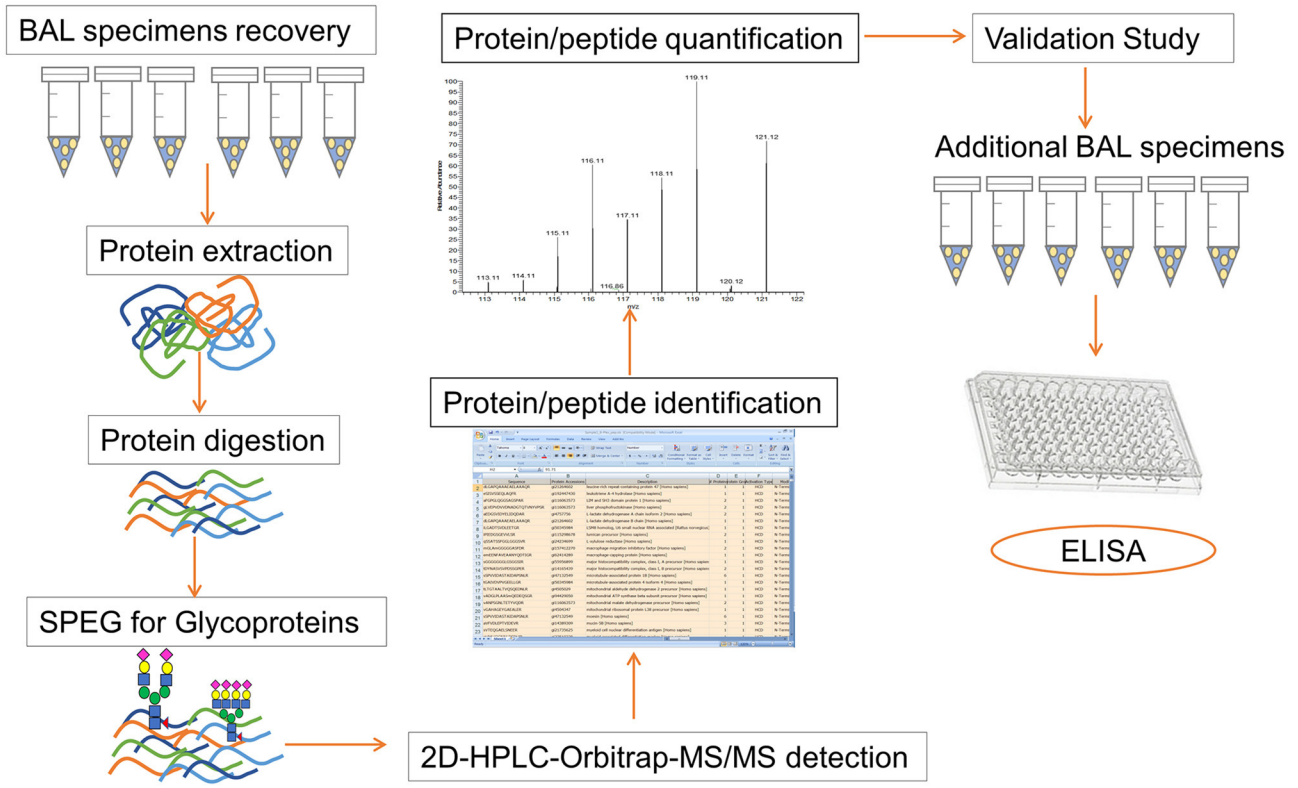


FIGURE 1 | Workflow for the proteomics detection in BAL specimen and subsequent validation study. The overall strategy of our study consists of following steps: (1) recovery of peptides from BAL samples, (2) capture of formerly *N*-linked glycopeptides using SPEG method, (3) Orbitrap Velos LC-MS/MS analysis, (4) identification and quantification of protein/peptide using Proteome Discoverer and SIEVE software, (5) analysis and comparison of the *N*-glycoprotein abundance in different BAL specimens, and (6) validation studies by using an ELISA assay.

the proteins were reduced by Tris (2-carboxyethyl) phosphine (TCEP) (Pierce, Rockford, IL) at room temperature for 1 h. BAL specimens were incubated with the sequencing grade trypsin (Promega, Madison, WI) at 37°C overnight. The digested peptides were used for further analysis.

Capture of *N*-Glycopeptides

As previously described, the SPEG method was used to capture *N*-linked glycopeptides (16). C18 column (Waters, Sep-Pak Vac) was used for peptide desalting, and tryptic peptides were eluted using 80% of acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA), oxidized with sodium periodate at a final concentration of 10 mM, and incubated for 1 h in the dark at room temperature. The oxidation reaction was quenched using sodium sulfite at a final concentration of 60 mM for 10 min at room temperature. The samples were coupled to hydrazide resin (Bio-Rad, Hercules, CA) in 80% ACN and 0.1% TFA for 4 h at room temperature. Then the resin was washing three times by using 100 mM NH_4HCO_3, 1.5 M NaCl, and H_2O to remove non-glycosylated peptides. *N*-linked glycopeptides were released from the resin by mixing with 1 μl PNGase F (New England Biolabs, Ipswich, MA) in 100 mM NH_4HCO_3 at 37°C overnight. After the C18 cartridge, the purified glycopeptides were dried by Speed-Vac and resuspended with 0.4% acetic acid for further use of LC-MS/MS analysis.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

Peptides were analyzed by LC-MS/MS on an Orbitrap Velos (Thermo Fisher Scientific, USA) and a nanoLC system (Eksigent, Dublin, USA). One microgram of peptides was subjected to a self-packed C18 column with a 120 min gradient. Mobile phase flow rate was set as 300 nL/min with 0.2% FA (A) and 0.2% FA, 95% ACN (B), respectively. The de-glycosylated peptides were separated with gradient elution of phase B from 5%, and ramped to 30% in 86 min. LC-MS/MS data acquisition used a data-dependent analysis (DDA) method of the top 10 most abundant ions. The MS spectra were collected from 350 to 1,800 *m/z* at a resolution of 70,000 followed by data-dependent higher-energy collisional dissociation (HCD) MS/MS, with a 2 *m/z* isolated mass window.

Protein Identification and Label-Free Quantification

Proteome Discoverer software (Thermo Fisher Scientific, USA) was applied for protein annotation. The dataset was searched

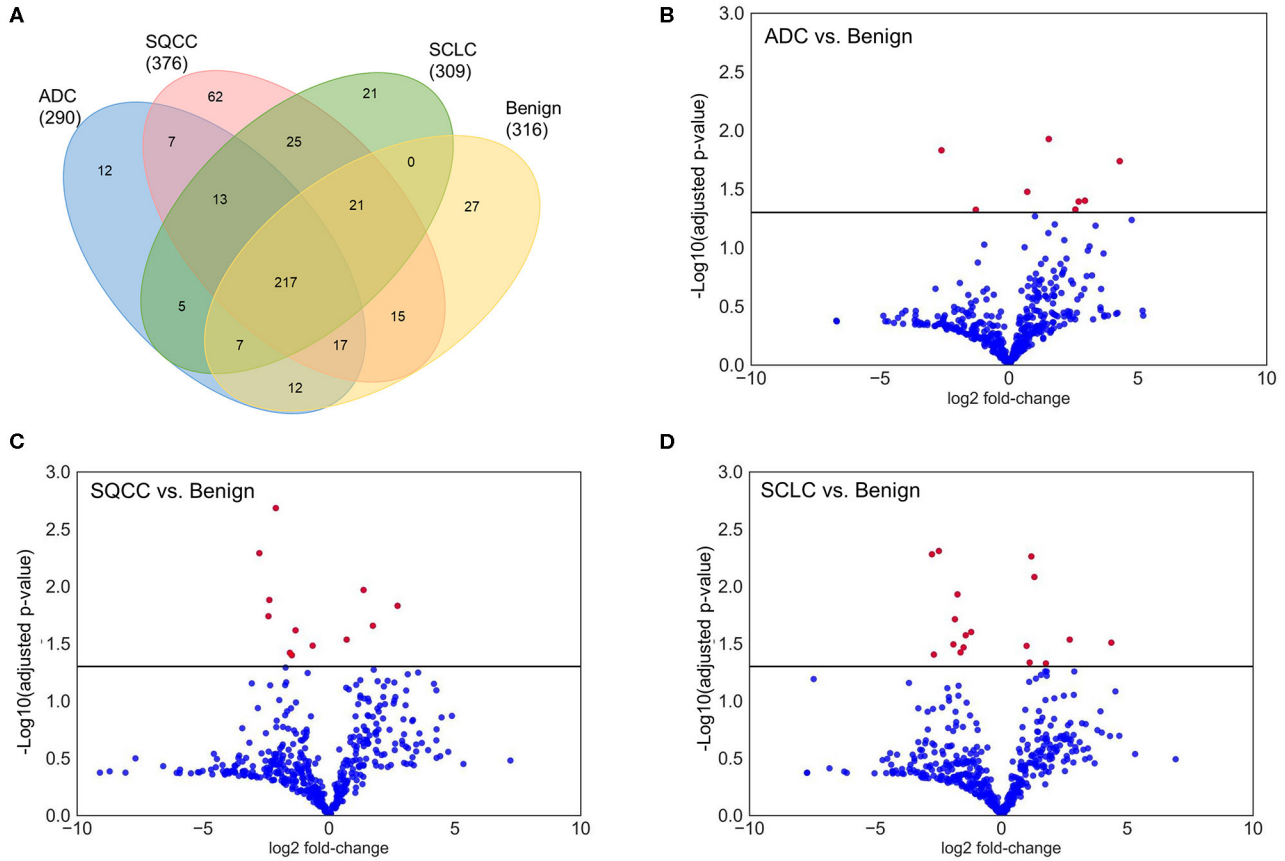


FIGURE 2 | Overview identification and quantification of glycoproteins in BAL samples. **(A)** Venn diagram of identified glycoproteins in lung ADCs, SQCCs, SCCs, and benign samples. **(B)** Volcano plot showing the variance in protein expression between lung ADCs and benign samples. **(C)** Volcano plot showing the variance in protein expression between lung SQCCs and benign samples. **(D)** Volcano plot showing the variance in protein expression between lung SCLCs and benign samples (Wilcoxon signed-rank test, Red dot: adjusted $P < 0.05$, Blue dot: adjusted $P \geq 0.05$).

against a human protein database (the International Protein Index human protein database containing a total number of 91,464 entries, version 3.87). The MS and MS/MS tolerances were set as 20 ppm and 0.06 Da, respectively. For de-glycosylated peptides identification, the cysteine residues were modified by carbamidomethyl group (C, +57.105 Da) as a stable modification, and methionine oxidation (M, +15.9949 Da) and asparagine deamidation (N, +0.98 Da) were set as dynamic modifications. A maximum of 2 miss cleavages was allowed. Then we used SIEVE software (Thermo Scientific, USA) for the label-free, semi-quantitative analysis of glycopeptides and glycoproteins. The chromatograms of raw files were aligned after importing into the SIEVE. Frame parameters were set as follows: mass range from 350 to 1,800, retention time from 0 to 120 min, a maximum frame of 100,000, frame *m/z* width of 10 ppm, frame retention time width for 2.5 min, the intensity threshold of 5×10^5. After the SIEVE framing, the results from Proteome Discoverer were filtered with 1% false-discovery rate (FDR) and quantified using the unique peptide for a protein. In SIEVE analysis, the frames for formal glycopeptides (i.e., those containing NXS/T, X≠ P, and N was deamidated) were targeted for quantification. Data were normalized by total ion current (TIC).

Enzyme-Linked Immunosorbent Assay (ELISA)

In the assay, 100 μl of BAL specimens were incubated with periostin capture antibody in a 96-well-microtiter plate overnight at 4°C, and then incubated with HRP-conjugated mouse anti-human periostin antibody at 4°C for 30 min. Tetramethylbenzidine was added and interacted at room temperature for 20 min. Then using an absorbance reader (Biotek, Winooski, VT) to acquire the absorbance readings at 450 nm. Recombinant human periostin (concentration range: 0–250 ng/mL) was used to generate a standard curve from a four-parameter logistic curve fit. All samples were analyzed duplicated and calculated the average of the two values. The total protein concentration of all BAL samples was determined using the BCA protein assay. The levels of periostin were expressed as ng/mg total proteins in BAL specimens.

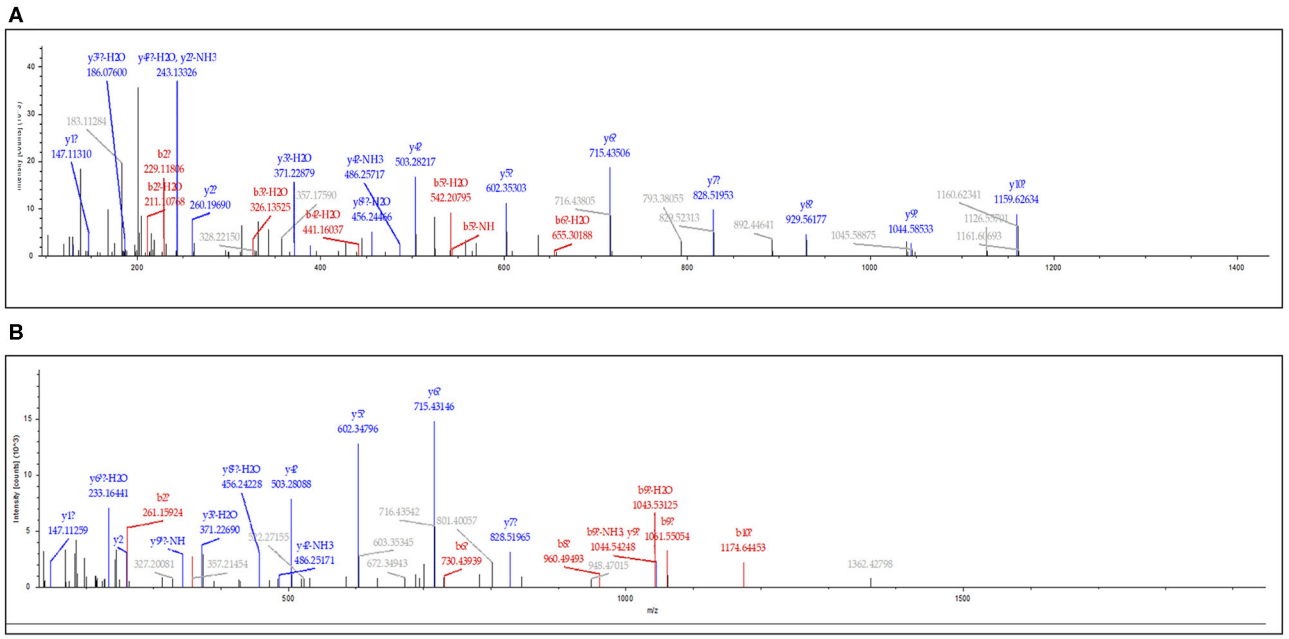


FIGURE 3 | Identification of periostin in lung BAL samples. The MS/MS spectrum corresponding to periostin peptide sequence EVN*DTLLVNELK (Charge: +2, Monoisotopic m/z: 694.37050 Da, MH+: 1,387.73373 Da) **(A)** and IFLKEVN*DTLLVNELK (Charge: +3, Monoisotopic m/z: 630.35892 Da, MH+: 1,889.06221 Da) **(B)**. *N** is the formerly *N*-linked glycosylation site and deaminated to Aspartic Acid (D) when *N*-glycans were released from the site by PNGase F.

Statistical Analysis

Characteristics of clinicopathological information were evaluated using the SPSS statistical software (version 24.0, SPSS Inc., USA) The χ^2 and Fisher's exact tests were used to calculate the *P*-value. The relationship between periostin and pathological characteristics of NSCLS patients was using Spearman rank correlation analysis. All tests were determined using bilateral 95% confidence intervals (CI). If the alpha value was <0.05, it was considered statistically significant ($P < 0.05$).

RESULTS

Clinical Information

Discarded BAL fluids were collected from the cytologic laboratory for the proteomic analysis, including ADC ($n = 5$), SQCC ($n = 4$), SCLC ($n = 4$), and benign lung diseases ($n = 3$). All cancer cases were primary lung cancers. For 16 proteomic BAL samples, the median age of patients was 63 years, ranging from 23 to 88 years; and the female and male ratio was: 1:1. In 13 cancer cases, 1 case was pT1, 3 cases were pT2, 5 cases were pT3, and 4 cases were pT4. The smoking status revealed that six patients were non-smokers, and 10 were former or current smokers (**Supplementary Table 1**).

Additional BAL specimens were collected for the ELISA assay, including six cases of benign lung diseases, 18 cases of primary lung ADCs, 6 cases of primary lung SCLCs, and 9 cases of primary lung SQCCs. Of 39 cases, the median age of patients was 65 years, ranging from 28 to 85 years. The female and male ratio was: 1.3:1. In 33 cancer cases, 4 cases were pT1, 10 cases were pT2, 12 cases were pT3 and 7 cases were pT4. The smoking status revealed that eight patients were non-smokers, and 31 were former or current smokers (**Supplementary Table 2**).

Detection of Glycoproteins and Periostin in BAL Specimens by Label-Free Quantification and SIEVE Analysis

The overall strategy in our study included the following main steps: (1) recovery of peptides from BAL samples, (2) capture of formerly *N*-linked glycopeptides using the SPEG method, (3) analysis of the *N*-glycoprotein abundance in different BAL specimens, and (4) validation of our findings using an ELISA assay (**Figure 1**).

In total, 462 glycoproteins from 1,081 glycopeptides were identified in BAL samples, whereas, there were 290, 376, 309, and 316 glycoproteins identified in ADCs, SQCCs, SCLCs, and benign samples, respectively. Compared to the benign samples, there were 8, 12, and 17 glycoproteins differentially expressed in ADCs, SQCCs, and SCLCs, respectively (adjusted $P < 0.05$, Wilcoxon signed-rank test). The differentially expressed glycoproteins included periostin (POSTN), vascular cell adhesion protein 1 (VCAM1), TNF Receptor Superfamily Member 19 (TNFRSF19), and mucin-13 (MUC13). Overall data are summarized in **Figure 2** and **Supplementary Table 3**.

Among these differentially expressed glycoproteins in lung cancer, periostin was a particularly interesting one. The periostin was identified in the BAL samples using a combination of *N*-glycopeptides SPEG capture with SIEVE label-free quantification. Two glycopeptides of periostin were detected with the sequences

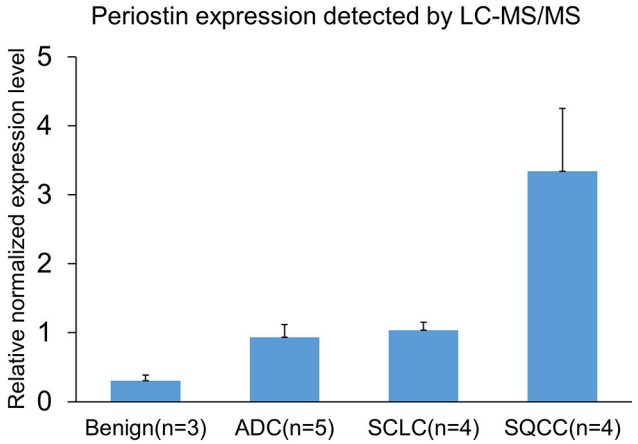


FIGURE 4 | Detection and quantification of periostin in BALs by mass-spectrometry (MS) and SIEVE software. Data are shown as relatively normalized expression levels (means ± SE). The mean values of periostin in benign disease, ADCs, SQCCs, and SCLCs were 0.30 ± 0.08, 0.94 ± 0.18, 3.34 ± 0.91, and 1.04 ± 0.12.

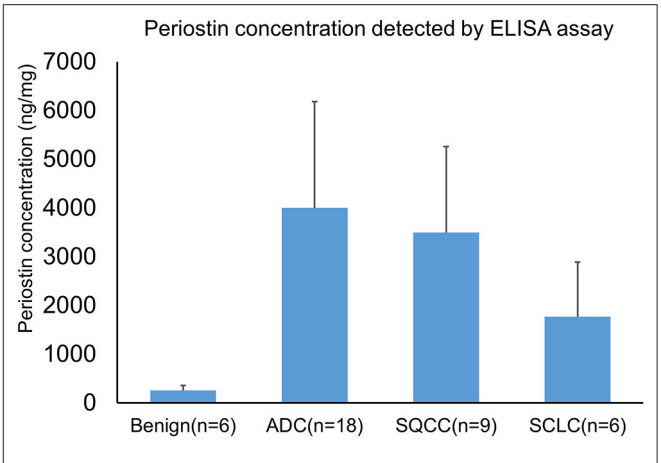


FIGURE 5 | Detection of periostin in BAL specimens by an ELISA assay. BAL specimens were independently collected from benign lung disease and lung cancer patients. A total of 39 BAL specimens were used for the assay. The protein concentration of each specimen was determined and used for the normalization of periostin levels. The mean expression levels of periostin in benign disease, ADCs, SQCCs, and SCLCs were 255 ± 104 (mean ± SE) and 4,002 ± 2,181, 3,496 ± 1,765, and 1,772 ± 1,119 ng/mg of total BAL proteins.

of EVN*DTLLVNELK and IFLKEVN*DTLLVNELK (**Figure 3**). The mean values (the relative normalized expression level) of periostin in benign disease, ADCs, SQCCs, and SCLCs were 0.30 ± 0.08 (mean ± SE), 0.94 ± 0.18, 3.34 ± 0.91, and 1.04 ± 0.12 (**Figure 4**). By comparing to the benign samples, we detected higher expressions of periostin in all three tumor types.

Validation of Periostin Expression in BAL by an ELISA Assay

We further investigated the expression of periostin in BAL using an independently collected cohort of lung cancer patients by an ELISA assay. In the assay, we included six cases of benign lung diseases, 18 cases of primary lung ADCs, 9 cases of primary lung SQCCs, and 6 cases of lung SCLCs. The protein concentration of individual specimens was also determined and used for the normalization of periostin levels. The average expression levels of periostin in benign disease, ADCs, SQCCs, and SCLCs were 255 ± 104 (mean ± SE) and 4,002 ± 2,181, 3,496 ± 1,765, and 1,772 ± 1,119 ng/mg of total BAL proteins (**Figure 5**). The periostin levels elevated in all cancer groups compared to the benign lung lesion group. These findings were in consistent with the LC-MS/MS data. However, we also found differential expression trends among lung cancer types.

DISCUSSION

Glycoproteins play a critical role in the regulation of cellular biological functions as well as in the discovery of cancer-associated proteins (33, 34). Glycosylation, especially the *N*-linked glycosylation, is commonly identified on a large number of cellular proteins, including cell surface proteins, transmembrane proteins, and extracellular proteins secreted in body fluids (33, 35). Studies have shown that glycoproteins secreted in body fluids such as urine, serum, cerebrospinal fluid (CSF), and BAL can be analyzed to monitor the disease progression in benign lung disease. These characteristics make glycoproteins as an indispensable group of potential cancer-associated proteins in the realm of clinical and biological research of cancers. We studied the profile of glycoproteins in the air-way fluid using the solid-phase extraction of *N*-glycoprotein and mass spectrometry-based LC-MS/MS approach, and then, validated our findings by using additional independent collected BAL specimens and ELISA assay. This is the first study utilizing air-way fluid to characterize *N*-glycoproteins expression in different types of lung cancer patients, particularly as proof-of-principle how BAL samples can be used for the study of cancer-associated extracellular glycoproteins in lung cancers.

In our study, a total of 462 glycoproteins were identified in the BAL samples, including 290, 376, 309, and 316 glycoproteins identified in ADCs, SQCCs, SCLCs, and benign samples, respectively. Compared to the benign samples, there were 8, 12, and 17 glycoproteins differentially expressed in ADCs, SQCCs, and SCLCs, respectively, such as POSTN, VCAM1, TNFRSF19, and MUC13.

VCAM-1 (CD106) is a 90-kDa glycoprotein, and frequently expressed in endothelial cells, which is participated in the regulation of cellular adhesion and inflammation-related activities (36, 37). The expression of VCAM-1 is associated with the activation of inflammatory cytokines, such as TNFα and ROS, and is relevant to the progression of several immunological disorders and cancers (38, 39). MUC13 is a cell surface mucin glycoprotein and plays a critical role in cancer cell proliferation and apoptosis (40, 41). It is often over-expressed in various cancer types, such as colorectal (42), gastric (43), pancreatic (44), and ovarian cancers (45). Some studies have also indicated its role in tumorigenesis by activating Wnt and NF-κB signaling pathways (46, 47). TNFRSF19, also known as TROY, is the member of TNF receptor superfamily, and usually transduces

cytokine signals through specific adaptor proteins that bind to the intracellular domain (ICD). High expression of TNFRSF19 is related to the inferior prognosis in melanoma (48), glioblastoma (49), and colorectal cancers (50), and is reported promoting tumorigenesis via TGFβ signaling pathways in nasopharyngeal carcinoma (51). We identified the differential expression of these proteins in the lung cancer BALs compared to the benign controls, which provided new insights for the discovery of lung cancer-associated glycoproteins. However, the role of these proteins still needs to be further investigated in lung cancers.

In addition to VCAM-1, MUC13, and TNFRSF18, periostin is a particularly interesting one. Periostin is a 93 kDa matrix *N*-glycoprotein, that was first identified in osteoblast-like cells. It is present in many healthy tissues and involved in bone formation, cardiac development, wound healing, and other cellular functions (19–21). Periostin is also known to play a significant role in tumorigenesis via the promotion of EMT of carcinoma cells and the regulation of the integrin signaling pathways, which leads to subsequent activation of the AKT/PKB (protein kinase B) and the FAK (focal adhesion kinase) network (19–24). A recent study has also suggested that the expression of periostin in fibroblasts may be involved in the process of premetastatic niche formation in the lung tissue (52). In lung cancer, the expression of periostin has been linked to tumor invasive and metastasis (23, 26–30). Recent studies have also shown that periostin can be detected in the serum of NSCLC patients (29, 30), indicating the potential utility of periostin as a therapeutic target in lung cancer. In an animal study using periostin knock-out mice, it has been demonstrated that the expression of periostin is involved in the phosphorylation of ERK in the Akt and FAK pathways (28). Furthermore, periostin is an extracellular protein and should be accessible via air-way fluid as well as other extracellular fluid. Recent studies have also demonstrated that the expression of periostin can be detected in the pleural effusions of NSCLC patients (53).

Quantitative proteomic analysis of BAL fluid has been applied to study certain benign lung diseases such as cystic fibrosis, asthma, interstitial lung disease (14, 15) and idiopathic eosinophilic pneumonia (54), but the protein profile in different types of lung cancers, especially the *N*-linked glycoproteomics, has not been well-studied. We detected the glycoprotein expression in BAL from all subtypes of lung cancers, including both small cell and non-small cell lung cancers. More importantly, we found that periostin could be detected in air-way fluids by glycoproteomic analysis, which was validated using an independently collected cohort of BAL samples. Our finding is also in accordance with the observation of elevated serum levels of periostin in lung cancer (29). Taken together, our study suggests that secreted periostin from tumor cells can be detected in the air-way fluid, indicating that it may play a tumorigenic role in lung cancer development.

In this study, we first analyzed *N*-glycoproteins including periostin expressions in lung cancer BALs by using MS-based analysis, and then, further validated our findings by ELISA assay. The sequential steps of our analysis provide strong evidence that the air-way fluid can be used for the development of a clinical assay for the detection of cancer-associated glycoproteins in personalized medicine. In the MS-based analysis, we found that the periostin was differentially expressed among all lung cancer types, with a relatively higher level in SQCC using the *N*-glycoprotein capture technique. Although we found a similar result in the subsequent ELISA assay among lung cancer types, the expression of periostin revealed a relatively higher level in ADC. Our finding of variable expression of periostin in lung cancers may be due to several factors, such as different detection techniques used in the assay and the relatively small sample size in our study. The role of periostin in lung cancers needs to be further investigated.

Our study has several strengths, we have revealed that mass spectrometry-based *N*-linked glycoprotein analysis of air-way fluid could be used for the identification of cancer-associated glycoproteins in lung cancer, most notably that it is the first study to characterize the glycoproteins as well as periostin expression in different types of lung cancer patients using BAL samples. It may provide valuable information for lung cancer patients. However, our study also has limitations, such as the relatively small sample size and the large standard deviation during analysis, particularly in the ELISA assay. These limitations indicate that the sample collections procedure needs to be standardized prior to the proteomic analysis and the validation study. Further efforts to improve the detection of cancer-associated glycoproteins in the air-way fluid and large cohorts of validations are required for our future study.

In summary, our study utilizes an unique approach for the detection of glycoproteins in lung air-way fluids and demonstrates that various proteins, including periostin, are differentially expressed in BAL in lung cancer patients. Our findings provide new insight into the biological role of periostin in lung cancers. Although our findings indicate the possibility that periostin may be a potential prognostic marker in lung cancer patients, the further large-scale study is still needed to validate our findings. As current research continues to focus on the study of the mechanism and prognostic significance of periostin, these data may contribute to refine our knowledge of periostin as a cancer-associated glycoprotein and lead to a better understanding its role in cancer biology.

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 at: MassIVE database, https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp.
Accession: MSV000085469.

ETHICS STATEMENT

This study was reviewed by the Institutional Review Board of the Johns Hopkins Medical Institution. Ethical approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent was not required to participate in this

study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YZ, WY, HZ, and QL: conceptualization. YZ, MA, EG, and QL: project administration and methodology. YZ, WY, MA, and NH: data curation and software. YZ, WY, and HZ: formal analysis. MA and NH: validation. YZ and NH: writing—original draft. EG, DC, HZ, and QL: writing—review and editing. All authors read and approved the final manuscript.

FUNDING

This work was partially supported by Drs. Ji and Li Family Cancer Research Foundation (QL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01072/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* (2018) 68:7–30. doi: 10.3322/caac.21442
2. Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. Introduction to the 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *J Thorac Oncol.* (2015) 10:1240–2. doi: 10.1097/JTO.0000000000000663
3. Network CGAR. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* (2014) 511:543–50. doi: 10.1038/nature13385
4. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers. *Nature.* (2012) 489:519–25. doi: 10.1038/nature11404
5. Munfus-McCray D, Harada S, Adams C, Askin F, Clark D, Gabrielson E, et al. EGFR and KRAS mutations in metastatic lung adenocarcinomas. *Hum Pathol.* (2011) 42:1447–53. doi: 10.1016/j.humpath.2010.12.011
6. Munfus-McCray D, Cui M, Zhang Z, Gabrielson E, Askin F, Li QK. Comparison of EGFR and KRAS mutations in primary and unpaired metastatic lung adenocarcinoma with potential chemotherapy effect. *Hum Pathol.* (2013) 44:1286–92. doi: 10.1016/j.humpath.2012.10.016
7. Li QK, Singh A, Biswal S, Askin F, Gabrielson E. KEAP1 gene mutations and NRF2 activation are common in pulmonary papillary adenocarcinoma. *J Hum Genet.* (2011) 56:230–4. doi: 10.1038/jhg.2010.172
8. Allison DB, Lilo MT, Geddes S, Pallavajjalla A, Askin F, Gabrielson E, et al. Detection of PIK3CA mutations, including a novel mutation of V344G in exon 4, in metastatic lung adenocarcinomas: a retrospective study of 115 FNA cases. *Cancer Cytopathol.* (2016) 124:485–92. doi: 10.1002/cncy.21714
9. Shaukat I, Kern JJ, Höti N, Zhang H, Li J, Zheng G, et al. Detection of RAS and RAS-associated alterations in primary lung adenocarcinomas. A correlation between molecular findings and tumor characteristics. *Hum Pathol.* (2019) 84:18–25. doi: 10.1016/j.humpath.2018.09.004
10. Stoll LM, Yung RCW, Clark DP, Li QK. Cytology of endobronchial ultrasound-guided transbronchial needle aspiration versus conventional transbronchial needle aspiration. *Cancer Cytopathol.* (2010) 118:278–86. doi: 10.1002/cncy.20103
11. Feller-Kopman D, Yung RCW, Burroughs F, Li QK. Cytology of endobronchial ultrasound-guided transbronchial needle aspiration: a retrospective study with histology correlation. *Cancer Cytopathol.* (2009) 117:482–90. doi: 10.1002/cncy.20049
12. Lilo MT, Allison DB, Younes BK, Cui M, Askin FB, Gabrielson E, et al. The critical role of EBUS-TBNA cytology in the staging of mediastinal lymph nodes in lung cancer patients: a correlation study with positron emission tomography findings. *Cancer Cytopathol.* (2017) 125:717–25. doi: 10.1002/cncy.21886
13. Zhou Y, Gong G, Wang H, Habibabady ZA, Lang P, Hales R, et al. Transthoracic fine-needle aspiration diagnosis of solid, subsolid, and partially calcified lung nodules: a retrospective study from a single academic center. *CytoJournal.* (2019) 16:16. doi: 10.4103/cytojournal.cytojournal_43_18
14. Magi B, Bargagli E, Bini L, Rottoli P. Proteome analysis of bronchoalveolar lavage in lung diseases. *Proteomics.* (2006) 6:6354–69. doi: 10.1002/pmic.200600303
15. Li QK, Gabrielson E, Askin F, Chan DW, Zhang H. Glycoproteomics using fluid-based specimens in the discovery of lung cancer protein biomarkers: promise and challenge. *Proteomics Clin Appl.* (2013) 7:55–69. doi: 10.1002/prca.201200105
16. Li QK, Shah P, Li Y, Aiyetan PO, Chen J, Yung R, et al. Glycoproteomic analysis of bronchoalveolar lavage (BAL) fluid identifies tumor-associated glycoproteins from lung adenocarcinoma. *J Proteome Res.* (2013) 12:3689–96. doi: 10.1021/pr400274w
17. Yang S, Chen L, Chan DW, Li QK, Zhang H. Protein signatures of molecular pathways in non-small cell lung carcinoma (NSCLC): comparison of glycoproteomics and global proteomics. *Clin Proteomics.* (2017) 14:31. doi: 10.1186/s12014-017-9166-9
18. Carvalho AS, Cuco CM, Lavareda C, Miguel F, Ventura M, Almeida S, et al. Bronchoalveolar lavage proteomics in patients with suspected lung cancer. *Sci Rep.* (2017) 7:42190. doi: 10.1038/srep42190
19. Bhowmick NA, Neilson EG, Moses HL. Stromal fibroblasts in cancer initiation and progression. *Nature.* (2004) 432:332–7. doi: 10.1038/nature03096
20. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer.* (2009) 9:265–73. doi: 10.1038/nrc2620
21. Ruan K, Bao S, Ouyang G. The multifaceted role of periostin in tumorigenesis. *Cell Mol Life Sci.* (2009) 66:2219–30. doi: 10.1007/s00018-009-0013-7
22. Bao S, Ouyang G, Bai X, Huang Z, Ma C, Liu M, et al. Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer Cell.* (2004) 5:329–39. doi: 10.1016/S1535-6108(04)00081-9
23. Hu W-W, Chen P-C, Chen J-M, Wu Y-M, Liu P-Y, Lu C-H, et al. Periostin promotes epithelial-mesenchymal transition via the MAPK/miR-381 axis in lung cancer. *Oncotarget.* (2017) 8:62248–60. doi: 10.18632/oncotarget.19273
24. Yu B, Wu K, Wang X, Zhang J, Wang L, Jiang Y, et al. Periostin secreted by cancer-associated fibroblasts promotes cancer stemness in head and neck cancer by activating protein tyrosine kinase 7. *Cell Death Dis.* (2018) 9:1082. doi: 10.1038/s41419-018-1116-6
25. Shao R, Bao S, Bai X, Blanchette C, Anderson RM, Dang T, et al. Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression. *Mol Cell Biol.* (2004) 24:3992–4003. doi: 10.1128/MCB.24.9.3992-4003.2004
26. Sakurai H, Maeshima A, Watanabe S, Suzuki K, Tsuchiya R, Maeshima AM, et al. Grade of stromal invasion in small adenocarcinoma of the lung: histopathological minimal invasion and prognosis. *Am J Surg Pathol.* (2004) 28:198–206. doi: 10.1097/00000478-200402000-00007
27. Nitsche U, Stangel D, Pan Z, Schlitter AM, Esposito I, Regel I, et al. Periostin and tumor-stroma interactions in non-small cell lung cancer. *Oncol Lett.* (2016) 12:3804–10. doi: 10.3892/ol.2016.5132
28. Okazaki T, Tamai K, Shibuya R, Nakamura M, Mochizuki M, Yamaguchi K, et al. Periostin is a negative prognostic factor and promotes cancer cell proliferation in non-small cell lung cancer. *Oncotarget.* (2018) 9:31187–99. doi: 10.18632/oncotarget.25435

29. Zhang Y, Yuan D, Yao Y, Sun W, Shi Y, Su X. Predictive and prognostic value of serum periostin in advanced non-small cell lung cancer patients receiving chemotherapy. *Tumor Biol.* (2017) 39:1010428317698367. doi: 10.1177/1010428317698367
30. Xu C-H, Wang W, Lin Y, Qian L-H, Zhang X-W, Wang Q-B, et al. Diagnostic and prognostic value of serum periostin in patients with non-small cell lung cancer. *Oncotarget.* (2017) 8:18746–53. doi: 10.18632/oncotarget.13004
31. Kudo Y, Siriwardena B, Hatano H, Ogawa I, Takata T. Periostin: novel diagnostic and therapeutic target for cancer. *Histol Histopathol.* (2007) 22:1167–74. doi: 10.14670/HH-22.1167
32. Tian Y, Choi CH, Li QK, Rahmatpanah FB, Chen X, Kim SR, et al. Overexpression of periostin in stroma positively associated with aggressive prostate cancer. *PLoS One.* (2015) 10:e0121502. doi: 10.1371/journal.pone.0121502
33. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell.* (2006) 126:855–67. doi: 10.1016/j.cell.2006.08.019
34. Zhou Y, Lih TM, Yang G, Chen SY, Chen L, Chan DW, et al. An integrated workflow for global, glyco-, and phospho-proteomic analysis of tumor tissues. *Anal Chem.* (2020) 92:1842–9. doi: 10.1021/acs.analchem.9b03753
35. Zhang H, Chan DW. Cancer biomarker discovery in plasma using a tissue-targeted proteomic approach. *Cancer Epidemiol Biomarkers Prev.* (2007) 16:1915–7. doi: 10.1158/1055-9965.EPI-07-0420
36. Rice GE, Bevilacqua MP. An inducible endothelial cell surface glycoprotein mediates melanoma adhesion. *Science (New York, NY).* (1989) 246:1303–6. doi: 10.1126/science.2588007
37. Osborn L, Hession C, Tizard R, Vassallo C, Luhowskyj S, Chi-Rosso G, et al. Direct expression cloning of vascular cell adhesion molecule 1, a cytokine-induced endothelial protein that binds to lymphocytes. *Cell.* (1989) 59:1203–11. doi: 10.1016/0092-8674(89)90775-7
38. Cook-Mills JM, Marchese ME, Abdala-Valencia H. Vascular cell adhesion molecule-1 expression and signaling during disease: regulation by reactive oxygen species and antioxidants. *Antioxidants Redox Signal.* (2011) 15:1607–38. doi: 10.1089/ars.2010.3522
39. Kong DH, Kim YK, Kim MR, Jang JH, Lee S. Emerging roles of vascular cell adhesion molecule-1 (VCAM-1) in immunological disorders and cancer. *Int J Mol Sci.* (2018) 19:1057. doi: 10.3390/ijms19041057
40. Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. Muc13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. *J Biol Chem.* (2001) 276:18327–36. doi: 10.1074/jbc.M008850200
41. Sheng YH, Lourie R, Linden SK, Jeffery PL, Roche D, Tran TV, et al. The MUC13 cell-surface mucin protects against intestinal inflammation by inhibiting epithelial cell apoptosis. *Gut.* (2011) 60:1661–70. doi: 10.1136/gut.2011.239194
42. Walsh MD, Young JP, Leggett BA, Williams SH, Jass JR, McGuckin MA. The MUC13 cell surface mucin is highly expressed by human colorectal carcinomas. *Hum Pathol.* (2007) 38:883–92. doi: 10.1016/j.humpath.2006.11.020
43. Shimamura T, Ito H, Shibahara J, Watanabe A, Hippo Y, Taniguchi H, et al. Overexpression of MUC13 is associated with intestinal-type gastric cancer. *Cancer Sci.* (2005) 96:265–73. doi: 10.1111/j.1349-7006.2005.00043.x
44. Khan S, Zafar N, Khan SS, Setua S, Behrman SW, Stiles ZE, et al. Clinical significance of MUC13 in pancreatic ductal adenocarcinoma. *HPB.* (2018) 20:563–72. doi: 10.1016/j.hpb.2017.12.003
45. Chauhan SC, Vannatta K, Ebeling MC, Vinayek N, Watanabe A, Pandey KK, et al. Expression and functions of transmembrane mucin MUC13 in ovarian cancer. *Cancer Res.* (2009) 69:765–74. doi: 10.1158/0008-5472.CAN-08-0587
46. Dai Y, Liu L, Zeng T, Liang JZ, Song Y, Chen K, et al. Overexpression of MUC13, a poor prognostic predictor, promotes cell growth by activating Wnt signaling in hepatocellular carcinoma. *Am J Pathol.* (2018) 188:378–91. doi: 10.1016/j.ajpath.2017.10.016
47. Sheng YH, He Y, Hasnain SZ, Wang R, Tong H, Clarke DT, et al. MUC13 protects colorectal cancer cells from death by activating the NF-κB pathway and is a potential therapeutic target. *Oncogene.* (2017) 36:700–13. doi: 10.1038/onc.2016.241
48. Spanjaard RA, Whren KM, Graves C, Bhawan J. Tumor necrosis factor receptor superfamily member TROY is a novel melanoma biomarker and potential therapeutic target. *Int J Cancer.* (2007) 120:1304–10. doi: 10.1002/ijc.22367
49. Loftus JC, Dhruv H, Tuncali S, Kloss J, Yang Z, Schumacher CA, et al. TROY (TNFRSF19) promotes glioblastoma survival signaling and therapeutic resistance. *Mol Cancer Res.* (2013) 11:865–74. doi: 10.1158/1541-7786.MCR-13-0008
50. Schon S, Flierman I, Ofner A, Stahringer A, Holdt LM, Kolligs FT, et al. β-catenin regulates NF-κB activity via TNFRSF19 in colorectal cancer cells. *Int J Cancer.* (2014) 135:1800–11. doi: 10.1002/ijc.28839
51. Deng C, Lin YX, Qi XK, He GP, Zhang Y, Zhang HJ, et al. TNFRSF19 inhibits TGFβ signaling through interaction with TGFβ receptor type I to promote tumorigenesis. *Cancer Res.* (2018) 78:3469–83. doi: 10.1158/0008-5472.CAN-17-3205
52. Kong J, Tian H, Zhang F, Zhang Z, Li J, Liu X, et al. Extracellular vesicles of carcinoma-associated fibroblasts creates a pre-metastatic niche in the lung through activating fibroblasts. *Mol Cancer.* (2019) 18:175. doi: 10.1186/s12943-019-1101-4
53. Wang J, Fu J, Shen Q, Zhang F, Wang Y, Wu LL. Identification and diagnostic value of pleural fluid periostin and serum periostin of malignant pleural effusions in patients with non-small-cell lung cancer. *J Clin Lab Anal.* (2019) 33:e22943. doi: 10.1002/jcla.22943
54. Katoh S, Matsumoto N, Tanaka H, Yasokawa N, Kittaka M, Kurose K, et al. Elevated levels of periostin and TGF-β1 in the bronchoalveolar lavage fluid of patients with idiopathic eosinophilic pneumonia. *Asian Pac J Allergy Immunol.* (2019). doi: 10.12932/AP-111018-0414. [Epub ahead of print].

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 © 2020 Zhou, Yang, Ao, Höti, Gabrielson, Chan, Zhang 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 Oncology

ORIGINAL RESEARCH
published: 24 July 2020
doi: 10.3389/fonc.2020.01032

Biliary Tract Carcinogenesis Model Based on Bile Metaproteomics

*Ariel A. Arteta [1,2,3], Miryan Sánchez-Jiménez [4], Diego F. Dávila [5], Oscar G. Palacios [5] and Nora Cardona-Castro [1,2,4]**

[1] School of Graduate Studies, CES University, Medellín, Colombia, [2] Basic Science Research Group, School of Medicine, CES University, Medellín, Colombia, [3] Associated Professor Department of Pathology, University of Antioquia, Medellín, Colombia, [4] Colombian Institute of Tropical Medicine (ICMT), Sabaneta, Colombia, [5] Department of Hepatobiliary and Pancreatic Surgery, CES Clinic, Medellín, Colombia

Purpose: To analyze human and bacteria proteomic profiles in bile, exposed to a tumor vs. non-tumor microenvironment, in order to identify differences between these conditions, which may contribute to a better understanding of pancreatic carcinogenesis.

Patients and Methods: Using liquid chromatography and mass spectrometry, human and bacterial proteomic profiles of a total of 20 bile samples (7 from gallstone (GS) patients, and 13 from pancreatic head ductal adenocarcinoma (PDAC) patients) that were collected during surgery and taken directly from the gallbladder, were compared. g:Profiler and KEGG (Kyoto Encyclopedia of Genes and Genomes) Mapper Reconstruct Pathway were used as the main comparative platform focusing on over-represented biological pathways among human proteins and interaction pathways among bacterial proteins.

Results: Three bacterial infection pathways were over-represented in the human PDAC group of proteins. IL-8 is the only human protein that coincides in the three pathways and this protein is only present in the PDAC group. Quantitative and qualitative differences in bacterial proteins suggest a dysbiotic microenvironment in the PDAC group, supported by significant participation of antibiotic biosynthesis enzymes. Prokaryotes interaction signaling pathways highlight the presence of zeatin in the GS group and surfactin in the PDAC group, the former in the metabolism of terpenoids and polyketides, and the latter in both metabolisms of terpenoids, polyketides and quorum sensing. Based on our findings, we propose a bacterial-induced carcinogenesis model for the biliary tract.

Conclusion: To the best of our knowledge this is the first study with the aim of comparing human and bacterial bile proteins in a tumor vs. non-tumor microenvironment. We proposed a new carcinogenesis model for the biliary tract based on bile metaproteomic findings. Our results suggest that bacteria may be key players in biliary tract carcinogenesis, in a long-lasting dysbiotic and epithelially harmful microenvironment, in which specific bacterial species' biofilm formation is of utmost importance. Our finding should be further explored in future using *in vitro* and *in vivo* investigations.

Keywords: pancreatic cancer, metaproteomic, proteomic, bile, zeatin, surfactin, IL-8, carcinogenesis model

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Alessandro Carrer, Veneto Institute of Molecular Medicine (VIMM), Italy
Daniele Vergara, University of Salento, Italy

***Correspondence:**
Nora Cardona-Castro
ncardona@ces.edu.co

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *05 February 2020*
Accepted: *26 May 2020*
Published: *24 July 2020*

Citation:
Arteta AA, Sánchez-Jiménez M, Dávila DF, Palacios OG and Cardona-Castro N (2020) Biliary Tract Carcinogenesis Model Based on Bile Metaproteomics. Front. Oncol. 10:1032. doi: 10.3389/fonc.2020.01032

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is a malignant and highly lethal neoplasm of unknown etiology and is usually diagnosed at advanced stages (1). The currently available surgical interventions and chemotherapeutic regimes are unable to provide the desired impact on disease outcomes, and there is a clear, dismal prognosis, as 70–80% of patients will succumb to this disease during the first 2 years post-diagnosis (2). PDAC is the fourth leading cause of cancer-related deaths worldwide (3) and is expected to become the second leading cause of cancer-related deaths by 2025, due both to the improved outcome of other malignancies, and on the stagnation in outcome improvement for PDAC over the past 30 years (4–6). Modifiable and non-modifiable risk factors for PDAC have an unconvincing molecular association with the disease. Modifiable factors seem to distribute haphazardly around the world, and the classic ones, such as tobacco, diabetes, gallstones (GS) and alcohol intake, are absent in a significant proportion of patients (7, 8). The development of interventions that successfully reduce the incidence of this lethal malignancy and improve its outcome is limited by the scarce knowledge of the molecular factors that may play a role in the complex process of PDAC carcinogenesis (9). Hence, any effort to better understand PDAC carcinogenesis, or to unravel novel therapies, may be the starting point in driving future clinical interventions.

Bacteria have been associated with benign and malignant disease, and bacterial carcinogenesis is a process still being characterized in detail. The knowledge from such study may be the starting point to drive clinical interventions focused on cancer prevention. The carcinogenesis associated with viruses is based on the integration of the viral genome into the host DNA (i.e., Human Papilloma Virus, Epstein-Barr) and has been extensively studied and characterized (10). Conversely, bacterial carcinogenesis is a phenomenon thought to be the result of epithelial cells' chronic exposure to a pro-inflammatory milieu exacerbated by bacteria (11, 12). However, this pro-inflammatory, physiopathological mechanism cannot explain convincingly by itself the development of carcinomas in the gastrointestinal and biliary tract, as inflammatory phenomena regularly occur throughout the human lifespan, and just a few human beings develop malignant neoplasms.

The biliary tract including intra-pancreatic bile ducts, is a semi-closed duct system possessing its own microbiota (13–15), lined by cholangiocytes, and in constant contact with bile. Cholangiocytes or cholangiocyte like cells are the proposed cell of origin for a range of biliary tract carcinomas, also named cholangiocarcinoma, in gallbladder, and intra or extrahepatic bile ducts (16). PDAC derives from ductal cholangiocytes or transdifferentiated acinar-to-ductal cholangiocytes (17, 18) covering intra-pancreatic bile ducts (ductal carcinoma), so from the histopathological point of view PDAC and biliary tract carcinomas are not very different (19). In the case of PDAC, local microbiota may have effects on oncogenesis (20) and long term survival (21), but most of the surveys associating PDAC and bacteria demonstrate spurious associations due to inconsistent isolation of specific bacterial species and the lack of a molecular basis for bacteria-induced carcinogenesis (22, 23). Being part of the biliary tract microenvironment, bacteria must contribute to bile protein pool composition in a similar way to cholangiocytes. As cholangiocytes and bacteria are in permanent contact with bile, we hypothesize that bile-associated protein changes could reflect bile duct system alterations in the microenvironment sufficient to transform benign epithelial cells into a malignant phenotype.

Bile is stored and concentrated in the gallbladder, which is a clean reservoir where this biological fluid can be extracted for protein analysis (24). In research, bile samples are typically taken from the distal portion of the biliary tract during endoscopic interventions, such as endoscopic retrograde cholangiopancreatography (ERCP) (25). However, the inflammatory process associated with biliary obstruction in most PDAC patients may alter bile protein composition in the distal portion of the biliary tract and limit the finding of meaningful biological information. The lack of meaningful biological findings hinders the development of a specific model of carcinogenesis for the biliary tract that takes into account its unique physiological conditions, and the interplay of human and bacterial proteins.

We analyzed samples of human bile taken directly from the gallbladder, and not by ERCP, exposed to a pancreatic tumor vs. non-tumor microenvironment. The aim of the study, once samples were analyzed by mass spectrometry, was to find meaningful biological information through pathway inference analysis of the identified human and bacterial proteins. Biological pathway analysis was initially performed using the g:Profiler platform to compare and generate a complete panorama of the gene-protein sets being analyzed, including over-represented KEGG biological pathways (26). Then, we directly analyzed each protein set using the KEGG Mapper Reconstruct Pathway module (27), focusing on over-represented pathways in g:Profiler for human proteins, and prokaryote interaction pathways for bacterial proteins. KEGG has become a world reference database for assisting biological interpretations of molecular data sets. Currently, biological pathway analyses are one of the most reliable strategies for mechanistic insights into omics data, since the kind of evidence that supports the statistical modeling is always experimental and manually curated (28). Thus, in this study, using a paradigm shifting metaproteomic approach, we aimed to unravel novel and meaningful biological information that may contribute to a better understanding of PDAC bacteria-induced carcinogenesis, proposing a new carcinogenesis model for the biliary tract.

MATERIALS AND METHODS

Ethics and Sample Acquisition

The Institutional Human Ethics Committee at CES University and Clinic approved this study, and patients must give informed consent. Samples were de-identified before performing proteomic analysis. A surgical pathologist collected a total of 20 gallbladder bile samples; seven from patients with gallstones (GS), and 13 from patients with PDAC arising from the head of the pancreas. All patients were Colombians, and

residents of Medellín (Colombia). For GS patients, bile was obtained in the operating room immediately after laparoscopic extraction of the gallbladder, puncturing the gallbladder fundus with a syringe, and aspirating at least 5 mL of bile. For PDAC patients, bile was similarly collected, by aspirating bile with a syringe from the gallbladder pancreatoduodenectomy specimens were sent to the pathology lab for a cryosection margin report. Immediately after collection, bile samples were transported on ice, aliquoted, and stored at −80°C until further analysis. Patients with a clinical history of previous malignant neoplasms, chemotherapy, prior biliary tract surgery or biliary stent placement, HIV, pregnancy, chronic pancreatitis, choledocholithiasis, cystic fibrosis, hepatolithiasis, primary biliary cholangitis, liver cirrhosis, primary sclerosing cholangitis, or acute cholecystitis were excluded from this study.

Protein Extraction

Bile samples were thawed at room temperature and processed as previously described with slight modifications (29). Briefly, 1 mL of bile was centrifuged for 10 min at 4°C and 3,000 rpm, and 1 mL of TRI reagent and 1 mL of chloroform were added. The mix was incubated for 5 min at room temperature (20–25°C) and centrifuged for 15 min at 4°C and 12.000 xg to separate proteins. Avoiding the central lipid layer, remaining tube contents (supernatant + pellet) were transferred to a new tube. Then, 1,200 μL of acetone was added, mixed, incubated for 4 h, and centrifuged for 15 min at 4°C at 12,000 xg. Acetone was discarded, and the tubes were dried at room temperature, after which 200 μL of reconstituting buffer was added to the pellet, and the solution dried and lyophilized.

Proteomic Analyses

Proteomic analysis was performed by Creative Proteomics (Ramsey Road, Shirley, NY 11967, USA), briefly, the techniques used are described as follows:

Sample Preparation for Proteomic Analysis

Total proteins were precipitated from the protein solution using methanol and chloroform. Approximately 10 μg of total protein was dissolved in 6 M urea aqueous solution and was denatured with 10 mM DL-dithiothreitol, incubated at 56°C for 1 h, followed by alkylation with 50 mM iodoacetamide, and incubated for 60 min at room temperature, protected from light. Next, 500 mM ammonium bicarbonate (ABC) was added to the solution to make a final concentration of 50 mM ABC with a pH of 7.8. Promega Trypsin was added to the protein solution for digestion at 37°C for 15 h. The generated peptides were further purified with the C18 SPE column (Thermo Scientific) to remove salt. Samples were dried in a vacufuge and stored at −20°C until use.

Nano Liquid Chromatography

An Easy-nLC1000 (ThermoFisher Scientific, USA) coupled to a 100 μm× 10 cm in-house made column packed with a reversed-phase ReproSil-Pur C18-AQ resin (3 μm, 120 Å, Dr. Maisch GmbH, Germany) was used. A sample volume of 5 μL was loaded, with a total flow rate of 600 nL/min, and a mobile phase of A: 0.1% formic acid in water; and B: 0.1% formic acid in acetonitrile. The analytical separation was run using a gradient: from 6 to 9% B for 15 min, from 9 to 14% B for 20 min, from 14 to 30% B for 60 min, from 30 to 40% B for 15 min and from 40 to 95% B for 3 min, eluting with 95% B for 7 min.

Mass Spectrometry and Data Analysis

An Orbitrap Q Exactive™ mass spectrometer (Thermo Fisher Scientific, USA) set on a spray voltage of 2.2 kV and a capillary temperature of 270°C was used. Mass spectrometry resolution was set to 70,000 at 400 m/z and precursor m/z range: between 300.0 and 1800.0. The production scan range starts from m/z 100, activated by collision-induced dissociation (CID), and an isolation width of 3.00. The raw files were analyzed and searched against the human protein database from Uniprot using Maxquant (1.5.6.5). The parameters were set as follows: the protein modifications were carbamidomethylation (C) (fixed), oxidation (M) (variable); the enzyme specificity was set to trypsin; the maximum missed cleavages was set to 2; the precursor ion mass tolerance was set to 10 ppm, and MS/MS tolerance was 0.6 Da.

Human and Bacteria Peptide-Protein List Selection for Analysis

Peptide-protein analysis was performed at ICMT-CES University. Contaminants, albumin, hemoglobin related peptides, and peptides with zero intensity were eliminated from the full human and bacteria list of peptides-proteins. The identifiers of protein were standardized, missing gene names were manually completed, and protein taxonomy was verified.

Then, the full list of shared proteins was adapted to meet the requirements of the Prostar platform online version 1.18.1 (30), seeking for differentially abundant human and bacterial proteins among groups (GS vs. PDAC). The intensity values were normalized with the mean centering method without including variance reduction. Partially observed values were imputed using the SLSA (Structured Least Squares Adaptive) method. The hypothesis test was performed using the Student's *t*-test, considering a logarithmic change of 2.5 and adjusting the false discovery rate to 0.42% (p-value = 0.00316). The biological validity of imputing non-existent values for non-observed proteins, in order to compare the exclusive groups of proteins, was explored. However we chose to perform the analysis based only on observed values in the two groups, GS and PDAC (shared proteins).

For further qualitative analysis, all the human protein lists of the total, exclusive and differentially abundant proteins from GS and PDAC patients (**Figure 1**) were included in the retrieve ID/mapping module of the Universal Protein consortium resource (Uniprot http://www.uniprot.org/, UniProt release 2019_10). Then, in order to provide mechanistic insights into the biologically integrated function, Uniprot-standardized human protein lists of entries for each group were analyzed in the g:Profiler web page (https://biit.cs.ut.ee/gprofiler/gost) (26). g:Profiler allows a multi-query approach, which performs an over-representative functional analysis of multiple protein-gene lists, comparing proteins among groups. Default options

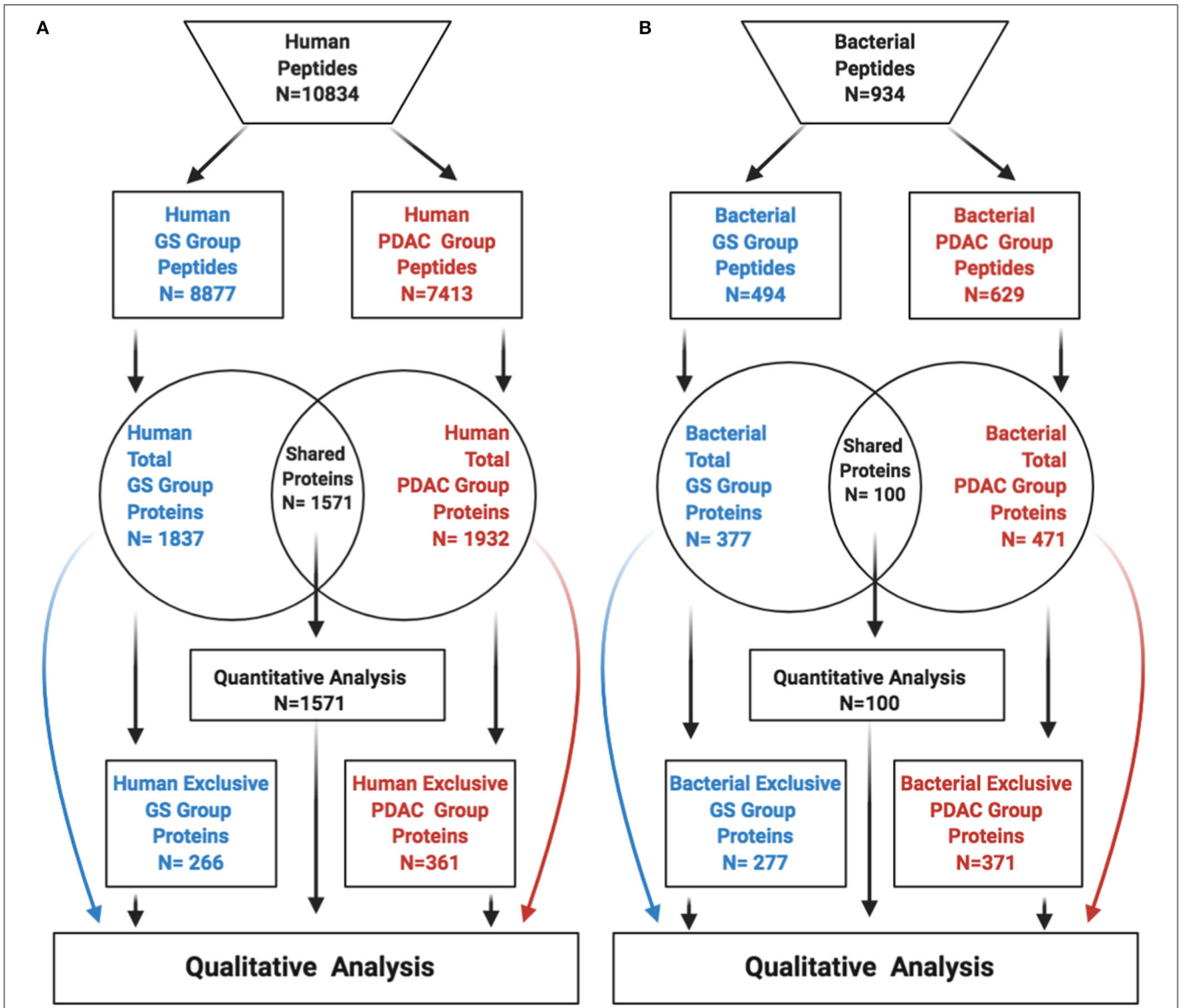


FIGURE 1 | Identified peptides-proteins per group and origin. The figure depicts peptides-proteins identified from human **(A)**, and bacteria **(B)**. GS, gallstones; PDAC, pancreatic ductal adenocarcinoma.

were maintained in g:Profiler, adding no electronic gene ontology annotations, and Bonferroni correction for multiple test adjustments. Significant, adjusted, over-represented pathways (p-values < 0.01), were used for further analysis in the KEGG Mapper Reconstruct Pathway. KEGG identifiers were obtained from the Uniprot FASTA file of the total, exclusive, and differentially abundant protein list, through BlastKoala (KEGG Orthology and Links Annotation version 2.2 https://www.kegg.jp/blastkoala/) (31). KEGG Mapper Reconstruct Pathway allows visualization and comparison of proteins in signaling pathways to identify qualitative and quantitative differences without coupled statistical analysis. (https://www.genome.jp/kegg/tool/map_pathway.html) (27). On the other hand, useful drugs were explored using the functional database DrugBank through WebGestalt (WEB-based GEne SeT AnaLysis Toolkit updated on 01/14/2019 http://www.webgestalt.org/) (32), by performing an over-representation analysis (ORA), using the database Drugbank and setting the false discovery rate at <0.01 (32) with Bonferroni correction for multiple test adjustments.

The biological context was analyzed as a whole for the total protein groups by correlating findings with the specific proteins identified for each condition. Many biological pathways were enriched over the Bonferroni p-adjusted value threshold in the total protein groups, but just three of them were related to bacterial infection (**Figure 2**).

For bacterial proteins, the same protocol for contaminant elimination, quality control, and differential abundance analysis was performed as for human proteins. We did not use g:Profiler

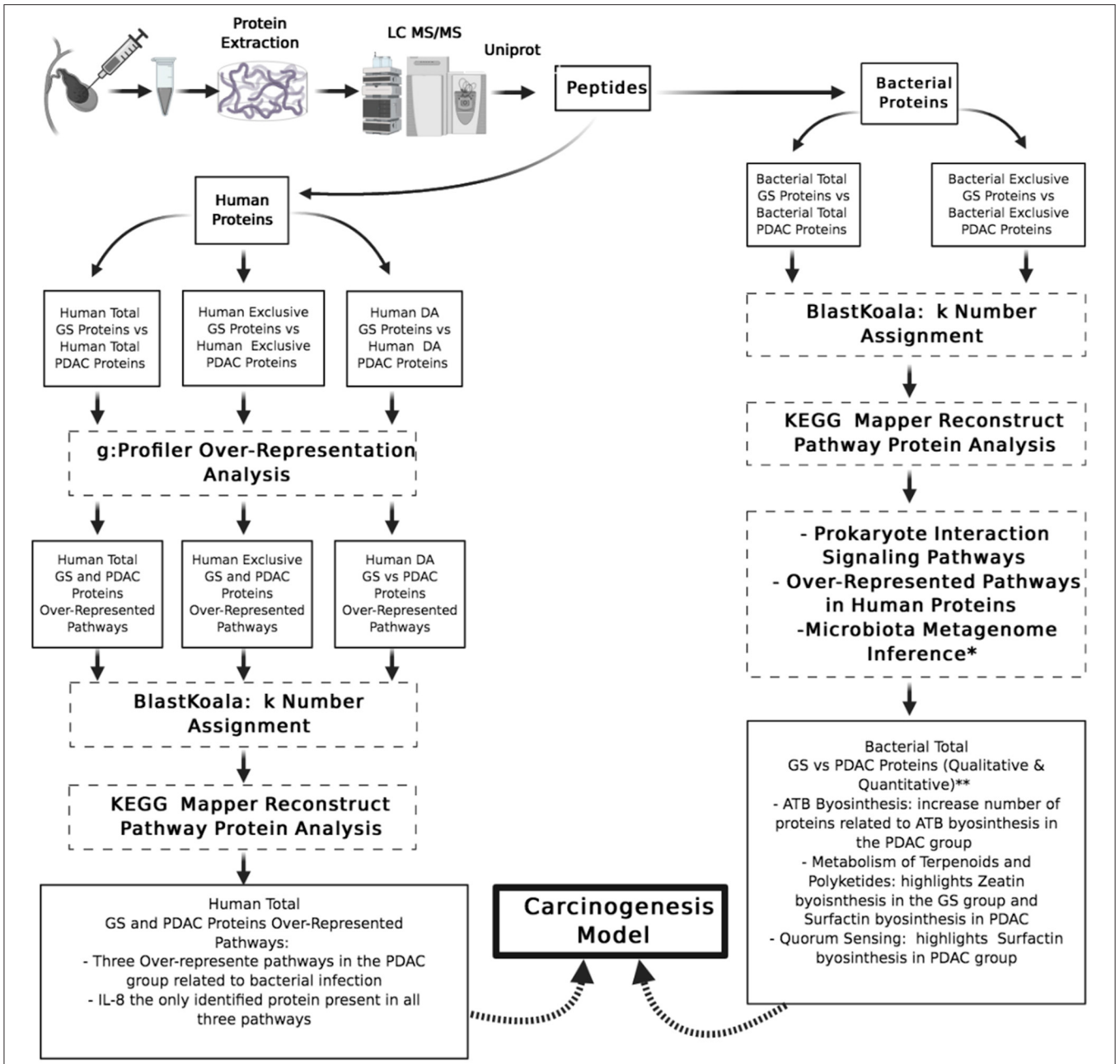


FIGURE 2 | Pipeline of proteins data sets analysis and summary or relevant findings. GS, gallstones; PDAC, pancreatic ductal adenocarcinoma; DA, differentially abundant; KEGG, Kyoto encyclopedia of genes and genomes; IL-8, interleukin 8; ATB, antibiotics. *Small sample microbiota metagenome inference analysis (unpublished results). **Not statistical analysis associated.

for bacterial protein analysis because this platform is not conceived for multi-species analysis. The bacterial protein lists of total and exclusive proteins from GS and PDAC patients were also included in the retrieve ID/mapping module of the Universal Protein consortium resource. KEGG identifiers were obtained using BlastKoala from the Uniprot FASTA files, and we focused our attention on prokaryote interaction signaling pathways during the analysis in KEGG Mapper Reconstruct Pathway (33).

RESULTS

A total of 20 bile samples extracted from gallbladders were analyzed, seven of which were taken from patients with GS (mean age of 48 years) (**Table 1**), and 13 from patients with PDAC (mean age of 56 years). All the patients were residents in Medellín and none of the patients presented clinical or histopathological signs of acute inflammation.

TABLE 1 | Clinical and demographic characteristics of patients.

Diagnosis	Age	Sex	Gallstones	Pathologic Staging	Perineural Invasion
GS	66	M	(+)	NA	NA
GS	52	F	(+)	NA	NA
GS	32	F	(+)	NA	NA
GS	42	M	(+)	NA	NA
GS	55	F	(+)	NA	NA
GS	48	M	(+)	NA	NA
GS	46	F	(+)	NA	NA
PDAC	52	F	(+)	pT2N0	(+)
PDAC	58	M	(+)	pT3N1	(+)
PDAC	66	M	(-)	pT2N1	(+)
PDAC	46	M	(-)	pT3N1	(+)
PDAC	68	M	(-)	pT3N1	(+)
PDAC	60	M	(-)	pT3N1	(+)
PDAC	35	M	(-)	pT2N1	(+)
PDAC	60	F	(+)	pT2N0	(+)
PDAC	56	F	(-)	pT3N0	(+)
PDAC	58	M	(-)	pT3N1	(+)
PDAC	61	F	(+)	pT2N1	(+)
PDAC	62	F	(-)	pT3N0	(+)
PDAC	55	M	(-)	pT3N1	(+)

GS, gallstones; PDAC, pancreatic ductal adenocarcinoma; M, male; F, female.

TABLE 2 | g:Profiler over-represented signaling pathway in human proteins.

Total Proteins GS = 1837 PDAC = 1932			
Term	**ID**	***p*-Adjusted value GS**	***p*-Adjusted value PDAC**
Pertussis	KEGG:05133	0.23886	0.00061
Proximal tubule bicarbonate reclamation	KEGG:04964	0.00094	0.06026
Cholesterol metabolism	KEGG:04979	0.03619	0.00105
Drug metabolism—other enzymes	KEGG:00983	0.01750	0.00111
Amino sugar and nucleotide sugar metabolism	KEGG:00520	0.00116	0.03220
Sulfur metabolism	KEGG:00920	0.00196	0.04153
Glyoxylate and dicarboxylate metabolism	KEGG:00630	0.03702	0.00203
Legionellosis	KEGG:05134	0.01740	0.00231
Adherens junction	KEGG:04520	0.00280	0.05119
Arginine and proline metabolism	KEGG:00330	0.02670	0.00305
Shigellosis	KEGG:05131	0.01019	0.00575
Exclusive Proteins GS = 266 PDAC = 361			
Phagosome	KEGG:04145	0.00011	0.04876
Metabolic pathways	KEGG:01100	1	0.00142
Differentially Abundant Proteins GS = 81 PDAC = 42			
Vasopressin-regulated water reabsorption	KEGG:04962	0.00066	1

KEGG, Kyoto Encyclopedia of Genes and Genomes; GS, gallstones; PDAC, pancreatic ductal adenocarcinoma.

After excluding peptides that were unassociated with any known proteins, a total of 10,834 human peptides were identified with a mean of 542 peptides per sample, 8,877 peptides in the GS group and 7,413 in the PDAC group. Peptides were associated with a total of 2,198 human proteins, 1,837 proteins in the GS group, and 1,932 proteins in the PDAC group. Upon comparison, a total of 1,571 proteins were common to both groups, while 266 proteins were exclusively found in the GS group, and 361 proteins in the PDAC group (**Figure 1A**). For bacteria, we identified a total of 934 peptides with a mean of 46 peptides per sample, 494 in the GS group and 629 in the PDAC group. Those peptides were associated with a total of 748 bacterial proteins, 377 proteins in the GS group and 471 in the PDAC group. We found 100 proteins shared among the two groups, with 277 exclusive proteins remaining in the GS group and 371 in the PDAC group (**Figure 1B**). Quantitative differential abundance analysis using Prostar revealed among the shared proteins within the human and bacteria groups, 123 differentially abundant human proteins, 81 in the GS group and 42 in the PDAC group, and no differentially abundant bacterial proteins.

Human Protein Over-Representation Analysis

The g:Profiler platform was used for the over-representation analysis in KEGG signaling pathways. Analyzing the total list of proteins, the platform identified from the 1,837 proteins in the GS group 1,832 (99.7%) and from 1,932 in the PDAC group 1,929 (99.8%). Regarding exclusive and differentially abundant proteins, the platform identified 100% of proteins in the GS and PDAC groups. In the total protein lists, we found five over-represented pathways in the GS group and seven in the PDAC group (**Table 2**), and in the exclusive protein lists, we identified one over-represented pathway in each group: phagosome in the GS group and metabolic pathways in the PDAC group. The analysis of the differentially abundant list of proteins revealed just one over-represented pathway in the GS group: vasopressin-regulated water reabsorption.

The over-represented pathways were analyzed in KEGG Mapper Reconstruct Pathway focusing our attention in the three g:Profiler over-represented pathways in the PDAC total protein group related to the bacterial infections Shigellosis, Pertussis, and Legionellosis. Analyzing the list of proteins in these three pathways (**Table 3**), it is notable that IL-8 (interleukin 8) is the only protein coinciding in the three pathways and present only in the PDAC group. This difference is more remarkable when evaluating the pathways of cytokine-cytokine receptor interaction and cytokines and growth factors, the latter in BRITE (Functional hierarchies of biological entities) tables, finding association in the presence of IL-8 with interleukin 11 (IL-11), CCL15 (Chemokine (C-C motif) ligand 15), CSF1 (Macrophage colony-stimulating factor) and CXCL7 (Chemokine (C-X-C motif) ligand 7) in the PDAC group (**Table 4**). Considering as an interaction point among prokaryotes and eukaryotes, Toll-like and NOD-like receptor

TABLE 3 | Human total proteins per group in KEGG of bacterial infection over-represented signaling pathways.

Shigellosis *N* = 20			Pertussis *N* = 22			Legionellosis *N* = 21		
Term	**Condition**	**Protein**	**Term**	**Condition**	**Protein**	**Term**	**Condition**	**Protein**
K04371	Gallstones:	MK01_HUMAN	K01330	Gallstones:	C1R_HUMAN	K01370	Gallstones:	
	Carcinoma:	MK03_HUMAN		Carcinoma:	C1R_HUMAN		Carcinoma:	CASP1_HUMAN
K04392	Gallstones:	RAC1_HUMAN	K01331	Gallstones:	C1S_HUMAN	K03231	Gallstones:	EF1A2_HUMAN
	Carcinoma:	RAC1_HUMAN		Carcinoma:	C1S_HUMAN		Carcinoma:	EF1A2_HUMAN
K04393	Gallstones:	CDC42_HUMAN	K01332	Gallstones:	CO2_HUMAN	K03233	Gallstones:	EF1G_HUMAN
	Carcinoma:	CDC42_HUMAN		Carcinoma:	CO2_HUMAN		Carcinoma:	EF1G_HUMAN
K04438	Gallstones:	CRK_HUMAN	K01370	Gallstones:		K03283	Gallstones:	HS71B_HUMAN
	Carcinoma:	CRK_HUMAN		Carcinoma:	CASP1_HUMAN		Carcinoma:	HS71B_HUMAN
K04514	Gallstones:		K02183	Gallstones:	CALM3_HUMAN	K03990	Gallstones:	CO3_HUMAN
	Carcinoma:	ROCK1_HUMAN		Carcinoma:	CALM3_HUMAN		Carcinoma:	CO3_HUMAN
K05692	Gallstones:	ACTB_HUMAN	K03986	Gallstones:	C1QA_HUMAN	K04077	Gallstones:	CH60_HUMAN
	Carcinoma:	ACTB_HUMAN		Carcinoma:	C1QA_HUMAN		Carcinoma:	CH60_HUMAN
K05700	Gallstones:	VINC_HUMAN	K03987	Gallstones:	C1QB_HUMAN	K04391	Gallstones:	CD14_HUMAN
	Carcinoma:	VINC_HUMAN		Carcinoma:	C1QB_HUMAN		Carcinoma:	CD14_HUMAN
K05748	Gallstones:	WASF2_HUMAN	K03988	Gallstones:	C1QC_HUMAN	K05482	Gallstones:	IL18_HUMAN
	Carcinoma:			Carcinoma:	C1QC_HUMAN		Carcinoma:	IL18_HUMAN
K05754	Gallstones:	ARPC5_HUMAN	K03989	Gallstones:	CO4A_HUMAN	K06461	Gallstones:	ITAM_HUMAN
	Carcinoma:	ARPC5_HUMAN		Carcinoma:	CO4A_HUMAN		Carcinoma:	ITAM_HUMAN
K05755	Gallstones:	ARPC4_HUMAN	K03990	Gallstones:	CO3_HUMAN	K06464	Gallstones:	
	Carcinoma:	ARPC4_HUMAN		Carcinoma:	CO3_HUMAN		Carcinoma:	ITB2_HUMAN
K05756	Gallstones:	ARPC3_HUMAN	K03994	Gallstones:	CO5_HUMAN	K07874	Gallstones:	RAB1A_HUMAN
	Carcinoma:	ARPC3_HUMAN		Carcinoma:	CO5_HUMAN		Carcinoma:	RAB1A_HUMAN
K05757	Gallstones:	ARC1B_HUMAN	K04001	Gallstones:	IC1_HUMAN	K07875	Gallstones:	RAB1B_HUMAN
	Carcinoma:	ARC1B_HUMAN		Carcinoma:	IC1_HUMAN		Carcinoma:	RAB1B_HUMAN
K05758	Gallstones:	ARPC2_HUMAN	K04002	Gallstones:	C4BPA_HUMAN	K07937	Gallstones:	ARF1_HUMAN
	Carcinoma:	ARPC2_HUMAN		Carcinoma:	C4BPA_HUMAN		Carcinoma:	ARF1_HUMAN
K05759	Gallstones:	PROF1_HUMAN	K04371	Gallstones:	MK01_HUMAN	K07953	Gallstones:	SAR1A_HUMAN
	Carcinoma:	PROF1_HUMAN		Carcinoma:	MK03_HUMAN		Carcinoma:	SAR1A_HUMAN
K06106	Gallstones:	SRC8_HUMAN	K04391	Gallstones:	CD14_HUMAN	K08517	Gallstones:	SC22B_HUMAN
	Carcinoma:	SRC8_HUMAN		Carcinoma:	CD14_HUMAN		Carcinoma:	SC22B_HUMAN
K06256	Gallstones:	CD44_HUMAN	K04513	Gallstones:	RHOA_HUMAN	K08738	Gallstones:	CYC_HUMAN
	Carcinoma:	CD44_HUMAN		Carcinoma:	RHOA_HUMAN		Carcinoma:	CYC_HUMAN
K07209	Gallstones:	IKKB_HUMAN	K04630	Gallstones:	GNAI2_HUMAN	K10030	Gallstones:	
	Carcinoma:			Carcinoma:	GNAI2_HUMAN		Carcinoma:	**IL8_HUMAN**
K07863	Gallstones:	RHOG_HUMAN	K05765	Gallstones:	COF1_HUMAN	K10159	Gallstones:	TLR2_HUMAN
	Carcinoma:	RHOG_HUMAN		Carcinoma:	COF1_HUMAN		Carcinoma:	
K10030	Gallstones:		K06461	Gallstones:	ITAM_HUMAN	K12799	Gallstones:	
	Carcinoma:	**IL8_HUMAN**		Carcinoma:	ITAM_HUMAN		Carcinoma:	ASC_HUMAN
K12836	Gallstones:	U2AF5_HUMAN	K06464	Gallstones:		K13525	Gallstones:	TERA_HUMAN
	Carcinoma:	U2AF5_HUMAN		Carcinoma:	ITB2_HUMAN		Carcinoma:	TERA_HUMAN
			K10030	Gallstones:		K15464	Gallstones:	BNIP3_HUMAN
				Carcinoma:	**IL8_HUMAN**		Carcinoma:	
			K12799	Gallstones:				
				Carcinoma:	ASC_HUMAN			

Kyoto Encyclopedia of Genes and Genomes. Bold indicates Interleukin-8.

pathways, we analyzed those signaling pathways and IL-8 was also present, and only in the PDAC group. In other KEGG signaling pathways with relevance to carcinogenesis processes such as DNA repair, xenobiotic metabolism, and pathways in cancer and pancreatic cancer, we didn't find differences. In the signaling pathways over-represented in exclusive and

TABLE 4 | Comparison of human total proteins related to cytokine-cytokine receptor interaction and growth factors and cytokines signaling pathways.

Cytokine-cytokine receptor interaction $N = 12$			Growth factors and cytokines $N = 8$		
Term	**Condition**	**Protein**	**Term**	**Condition**	**Protein**
K04723	Gallstones:	IL1AP_HUMAN	K05417	Gallstones:	
	Carcinoma:	IL1AP_HUMAN		Carcinoma:	IL11_HUMAN
K05059	Gallstones:	CNTFR_HUMAN	K05424	Gallstones:	LEP_HUMAN
	Carcinoma:			Carcinoma:	LEP_HUMAN
K05060	Gallstones:	IL6RB_HUMAN	K05450	Gallstones:	
	Carcinoma:	IL6RB_HUMAN		Carcinoma:	PDGFD_HUMAN
K05062	Gallstones:	LEPR_HUMAN	K05453	Gallstones:	
	Carcinoma:	LEPR_HUMAN		Carcinoma:	CSF1_HUMAN
K05090	Gallstones:	CSF1R_HUMAN	K05482	Gallstones:	IL18_HUMAN
	Carcinoma:			Carcinoma:	IL18_HUMAN
K05417	Gallstones:		K05511	Gallstones:	
	Carcinoma:	IL11_HUMAN		Carcinoma:	CCL15_HUMAN
K05424	Gallstones:	LEP_HUMAN	K10029	Gallstones:	
	Carcinoma:	LEP_HUMAN		Carcinoma:	CXCL7_HUMAN
K05453	Gallstones:		K10030	Gallstones:	
	Carcinoma:	CSF1_HUMAN		Carcinoma:	IL8_HUMAN
K05482	Gallstones:	IL18_HUMAN			
	Carcinoma:	IL18_HUMAN			
K05511	Gallstones:				
	Carcinoma:	CCL15_HUMAN			
K10029	Gallstones:				
	Carcinoma:	CXCL7_HUMAN			
K10030	Gallstones:				
	Carcinoma:	IL8_HUMAN			

Kyoto Encyclopedia of Genes and Genomes.

TABLE 5 | Average of the 5 most abundant taxonomic levels per group from identified total bacterial protein.

	Group			
	GS	**%**	**PDAC**	**%**
Phylum	Proteobacteria	60	Proteobacteria	64
	Firmicutes	15	Firmicutes	13
	Actinobacteria	7	Actinobacteria	6
	Cyanobacteria	4	Cyanobacteria	3
	Bacteroidetes	2,5	Thermotoga	2,8
Class	Gammaproteobacteria	32	Gammaproteobacteria	38
	Alfaproteobacteria	16	Alfaproteobacteria	17
	Bacilli	10	Bacilli	8
	Actinobacteria	7	Betaproteobacteria	6,5
	Deltaproteobacteria	4	Clostridia	6
Order	Enterobacterales	13	Enterobacterales	20
	Bacillales	7	Rhizobiales	6
	Rhizobiales	6,5	Clostridiales	5
	Clostridiales	5	Bacillales	4,5
	Pseudomonadales	4,5	Burkholderiales	4,5
Family	Enterobacteriaceae	8,5	Enterobacteriaceae	13
	Bacillaceae	4,5	Pasteurellaceae	3,5
	Pseudomonadaceae	3,5	Erwiniaceae	3,2
	Geobacteraceae	3	Bacillaceae	3
	Mycobacteriaceae	2,5	Burkholderiaceae	3
Genus	Escherichia	5	Shigella	5
	Bacillus	4	Escherichia	3
	Pseudomonas	3	Bacillus	2,5
	Geobacter	3	Salmonella	2,5
	Mycobacterium	2	Shewanella	2

GS, gallstones; PDAC, pancreatic ductal adenocarcinoma.

differentially abundant proteins, we could not find a clear biological meaning. In the WebGestal platform analysis, there were no relevant results for useful, therapeutic drugs using the Drugbank database.

Bacterial Protein Analysis

Proportional participation in some taxonomic levels of the imputed protein species is summarized in **Table 5**. The total protein list was analyzed in KEGG Mapper Reconstruct Pathway focusing on: (1) signaling pathways related to prokaryote interaction, (2) g:Profiler over-represented pathways in human proteins and (3) over-represented pathways in a metagenomic inference analysis (34). The latter analysis was performed from a small microbiota survey within the 20 samples, using bile from the gallbladders of GS patients ($N = 3$), bile from the gallbladders of PDAC patients ($N = 11$) and common biliary brush over the tumor from PDAC patients ($N = 11$) as samples. The results of the analysis show two statistically significant over-represented pathways, pyrimidine deoxyribonucleotide biosynthesis and isoprene biosynthesis (unpublished results).

Upon comparison of the GS and PDAC total protein groups, we found qualitative and quantitative differences regarding quorum sensing, biofilm formation, antibiotic synthesis (biosynthesis of other secondary metabolites) and metabolism of terpenoids and polyketides (**Table 6**). Regarding metabolism of terpenoids and polyketides, there is a protein involved in zeatin biosynthesis (MIAA_PSECP) that stands out from the other proteins, as it is not present in the PDAC group, is not a protein with an antibiotic function and is specific to that metabolic pathway. Concerning the PDAC group, in the signaling pathway of terpenoid and polyketide metabolism, there is one protein related to surfactin biosynthesis (SRFAB_BACSU), which is also notable, since this protein is also involved in the quorum-sensing signaling pathway. The proteins involved in quorum sensing and biofilm formation show qualitative differences, and the number of proteins present in antibiotic biosynthesis are considerably higher in the PDAC group compared to the GS group. The analysis of bacterial proteins present in the three g:Profiler over-represented human protein signaling pathways related to bacterial infection in the total PDAC group, show no differences in KEGG Mapper Reconstruct Pathway.

TABLE 6 | Comparison of bacterial list of total proteins in prokaryote interaction signaling pathways and microbiota metagenomic inference over-represented pathways.

Metabolism of terpenoids and polyketides				
Protein	**Gen**	**Phylum**	**Gram**	**Pathway**
GALLSTONES				
ISPE_RHOP5	ispE	Proteobacteria	Negative	Terpenoid backbone biosynthesis
ISPG_HAEIE	ispG	Proteobacteria	Negative	Terpenoid backbone biosynthesis
ISPG_SHEON	ispG	Proteobacteria	Negative	Terpenoid backbone biosynthesis
MIAA_PSECP	**miaA**	**Actinobacteria**	**Positive**	**Zeatin biosynthesis**
FADJ_PECCP	fadJ	Proteobacteria	Negative	Limonene and pinene degradation Geraniol degradation
TKT1_ECOLI	tktA	Proteobacteria	Negative	Biosynthesis of ansamycins
CARCINOMA				
ISPE_RHOFT	ispE	Proteobacteria	Negative	Terpenoid backbone biosynthesis
ISPG_GRABC	ispG	Proteobacteria	Negative	Terpenoid backbone biosynthesis
SRFAB_BACSU	**srfAB**	**Firmicutes**	**Positive**	**Nonribosomal peptide structures**
Biosynthesis of others secondary metabolites				
Protein	**Gen**	**Phylum**	**Gram**	**Pathway**
GALLSTONES				
DAPB_STRAW	dapB	Actinobacteria	Positive	Monobactam biosynthesis
CARCINOMA				
PROA_LEPBL	proA	Spirochaetes	Negative	Carbapenem biosynthesis
PROA_ALKOO	proA	Firmicutes	Positive	Carbapenem biosynthesis
DAPB_GEOKA	dapB	Firmicutes	Positive	Monobactam biosynthesis
DAPB_STRAW	dapB	Actinobacteria	Positive	Monobactam biosynthesis
DAPA_RHOFT	dapA	Proteobacteria	Negative	Monobactam biosynthesis
PGCA_BACSU	pgcA	Firmicutes	Positive	Streptomycin biosynthesis
HIS8_WOLSU	hisC	Proteobacteria	Negative	Novobiocin biosynthesis
TRPE_THET8	trpE	Deinococcus	Negative	Phenazine biosynthesis
Cellular Community—Prokaryotes				
Protein	**Gen**	**Phylum**	**Gram**	**Pathway**
GALLSTONES				
SECA_LACAC	secA	Firmicutes	Positive	Quorum sensing
SECA_CLOTH	secA	Firmicutes	Positive	Quorum sensing
YIDC_METCA	yidC	Proteobacteria	Negative	Quorum sensing
LUXS_DESPS	luxS	Proteobacteria	Negative	Quorum sensing, biofilm formation
CARCINOMA				
SP0A_CLOBU	spo0A	Firmicutes	Negative	Quorum sensing
SECA_CLOTH	secA	Firmicutes	Positive	Quorum sensing
SRFAB_BACSU	**srfAB**	**Firmicutes**	**Positive**	**Quorum sensing**
PTGA_SHIFL	crr	Proteobacteria	Negative	Biofilm formation
TRPE_THET8	trpE	Deinococcus	Negative	Biofilm formation
CHEB2_BURPS	cheB2	Proteobacteria	Negative	Biofilm formation

Source: Kyoto Encyclopedia of Genes and Genomes. Bold indicates Bacterial proteins included in the carcinogenesis model.

DISCUSSION

To the best of our knowledge, this is the first study to compare human and bacterial proteins, using a metaproteomic approach, bile samples exposed to a tumor vs. non-tumor environment in human PDAC and GS patients, respectively. The characterization of a single species protein profile is known as proteomics, while the characterization of a multi-species protein profile is

known as metaproteomics (35). The metaproteomic concept has been studied in humans through the characterization of fecal microbiota and the proteins produced by the different local bacterial species, enabling a better comprehension of the local conditions in the gastrointestinal tract (36, 37). In theory, the microenvironment within the biliary tract and the gallbladder will be more resistant to external variation and more accessible for bile retrieval in animal models. For all that, the biliary tract including its reservoir, will be an ideal biological system to evaluate through metaproteomic and microbiota analyses in conjunction, changes related to specific diets, neoplastic conditions, antibiotic use, chemotherapeutic schemas etc.

Finding meaningful biological information from omics' science data sets has been one of the major challenges of science in recent years (38). The relevance of research findings cannot be measured in every biological instance using statistical significance alone, as not all statistically significant results translate into meaningful biological change. Accordingly, in some areas of science, in which we cannot use statistics, or for which we have not developed appropriate tools, we should look for procedural alternatives that at least enable us to explore the real biological value of data sets. In our research, the three g:Profiler over-represented pathways in human proteins show qualitative and quantitative coincidences and differences regarding the presence of certain proteins in the PDAC and GS groups. The detailed analysis of the proteins in each over-represented signaling pathway is the component of the analysis with the greatest importance. Due to the polyfunctionality of bacteria and human proteins, these proteins must be contextualized and analyzed for relevant biological pathways.

IL-8: Carcinogenesis and Prokaryote Interactions

IL-8 was identified as a common protein in the three g:Profiler over-represented signaling pathways in PDAC human total proteins, associated with bacterial infections. IL-8 is a human chemotactic interleukin of the C-X-C family also known as CXCL8, originally discovered in macrophages (39), but also produced by epithelial cells. The effect of IL-8 depends on its interaction with specific membrane receptors coupled to G proteins CXCR1 (C-X-C motif chemokine receptor 1, C-X-C) and CXCR2 (C-X-C motif chemokine receptor 2) (40). Under physiological conditions IL-8 levels are undetectable, increasing in the presence of other pro-inflammatory cytokines like tumor necrosis factor α (TNFα) and interleukin 1β (41, 42). None of the latter two cytokines were identified in the PDAC or GS group, suggesting that alternative pathways can stimulate IL-8 synthesis.

High levels of IL-8 are described as poor outcome predictors in many malignant neoplasms, including PDAC. The cellular endpoint effects induced by the IL-8 CXCR1/CXCR2 axis, in normal epithelial cells, tumor cells or other cells in the tumor microenvironment, promote cellular survival, proliferation, angiogenesis, and a stem cell phenotype (43, 44). Concordantly, high levels of IL-8 in patients with breast, prostate and lung carcinoma, and melanoma are related to aggressive tumor behavior, due to high proliferation rate, local invasion, angiogenesis, and an increase of a stem cell phenotype and metastasis (45). In the special case of PDAC, high levels of IL-8 are also related to aggressive tumor behavior and poor prognosis, with evidence that includes PDAC cell line models (46), high blood levels in PDAC and cholangiocarcinoma patients (47, 48), and over-expression of IL-8 and its receptors in tumor tissue (49) and inflammatory cells infiltrating the tumor (50).

The biological relevance of IL-8 is not limited to neoplasms; there is a special prokaryote behavior linked to the synthesis of this interleukin. Biofilm formation by bacteria such as *F. nucleatum* and *A. naeslundii*, and not the planktonic form, stimulates the synthesis of IL-8 by human squamous epithelial cells (51). Supporting the latter concept, several surveys have proved that bacteria biofilm not only stimulates IL-8 synthesis by human squamous epithelial cells, but that stimulation is stronger when the biofilm is formed by multiple bacterial species (52, 53). Furthermore, the similarity of some amino acids in the carboxy-terminal region of IL-8 with cecropins, proteins with antibiotic properties, elicit the analysis of the antibiotic properties of IL-8 through the synthesis of synthetic peptides. These synthetic peptides are thought to be physiologically generated from acidic hydrolysis, and effectively have antibiotic properties which vary according to salt concentration and pH (54).

Some of the PDAC-specific proteins associated with IL-8 are also considered as poor outcome biomarkers in the natural history of malignant neoplasms. High levels of CXCL7 in cholangiocarcinoma tumoral tissue are associated with poor tumor differentiation, local lymph node metastasis, and lymphatic/vascular invasion (55). In renal carcinoma, high levels of CXCL7 are proposed as prognostic factors of chemotherapeutic response (56), and in colon cancer are related to poor survival in patients with liver metastasis (57). Similarly, high levels of CCL17 and IL-11 are associated with poor outcome in malignant neoplasms due to aggressive biological behavior regarding local invasion and metastasis (58–61).

Differences in Prokaryote Interaction Pathways

The metagenomic inference analysis results from the small microbiota group revealed some over-represented pathways. Of special interest is the metabolism of terpenoids and polyketides signaling pathway. This pathway was analyzed using the total list of bacterial proteins, finding qualitative and quantitative differences, among them the presence of zeatin in the GS group and surfactin in the PDAC group. Terpenoids and polyketides are a huge group of substances synthesized by bacteria, fungi, plants, and animals. Zeatin is an isoprenoid derived from adenine with two isoforms, trans and cis, depending on which of the two hydroxyl groups in the lateral chain of isopentenyl is hydroxylated (62). The identified bacterial protein in the GS group participates in the metabolic pathway for *cis*-zeatin synthesis and is specific to this metabolic pathway. Cis-zeatin is a cytokinin that can be produced by multiple bacterial species (63, 64), with just one published piece of research evaluating its activity in tumor cell lines, proving its anti-tumor potential in leukemia cell lines (65).

The ability to produce surfactin is a property of bacteria from the genus Bacillus, and since the discovery of surfactin in 1968 by Arima, this amphipathic lipopeptide has been found to possess several properties (66). Within these properties are those general to all lipopeptides and antibacterial proteins which act upon Gram-positive and Gram-negative bacteria (67, 68), and other anti-inflammatory (69) and anti-viral effects (70). Of interest in our research, the properties associated with biofilm formation are of utmost importance. Regarding biofilm formation, surfactin has a selective effect, primarily inhibitory, over many bacterial species, though to date there is no clear biological explanation for the selectivity of surfactin for biofilm production/inhibition (71–73).

Besides producing essential compounds for survival, bacteria are able to produce and secrete into the environment low molecular weight compounds called secondary metabolites. Within those secondary metabolites are substances with antibiotic properties that, in a specific microenvironment, confer an advantage upon the bacterium producing the antibiotics, reducing the number of competitors (mainly for nutrient acquisition) (74). Bacterial antibiotic synthesis is a phenomenon influenced by the community and denotes a competitive behavior for survival, and is seemingly species-specific (75). In our analysis, the increased number of identified proteins in the antibiotic synthesis pathways in the PDAC group compared to the GS group is remarkable. Based on that fact, we infer a major competition among species in the PDAC group.

The change in bacteria association, from free-living or planktonic to biofilm formation, relies upon many genetic factors and local conditions (76). Biofilm formation is tightly associated in multiple bacterial species with the increase of c-di-GMP (cyclic diguanylate) intracellular levels, which can also be induced by quorum sensing proteins (77). By means of conventional microbiota analysis, we are unable to determine if the identified bacteria are in a biofilm or not. For this reason, we considered that finding different bacterial proteins among the groups (PDAC and GS) related to quorum sensing and biofilm formation is of biological relevance, as these proteins can be involved in the process of bacteria-induced carcinogenesis. Inflammation is considered a starting point of bacteria associated carcinogenesis. However, this physio-pathological mechanism cannot fully explain the development of carcinomas in the gastrointestinal and biliary tract, as inflammatory states constantly occur throughout the human lifespan, and only a few human beings develop malignant neoplasms in the gastrointestinal system. The bacteria-cancer relationship has been viewed in a reductionist manner as simply a pro-inflammatory milieu initiated by bacteria, in line with the hypothesis of inflammation and cancer proposed by Virchow in 1835 (78). Previous studies have proposed that this pro-inflammatory milieu may be initiated by dysbiosis, which is defined as a change in the normal composition of the microbiota. However, dysbiosis has neither fulfilled the expectations nor provided—to date—a reliable molecular explanation for bacteria-induced carcinogenesis (79).

The Carcinogenesis Model

We hypothesized that there is no such thing as a dysbiotic microbiota, regarding the presence or absence of certain bacterial species. A dysbiotic microbiota is a haphazard composition of bacterial species with products harmful to bacteria and epithelial cells, specific to a particular individual vis-à-vis microbiota-modifying factors. Based on our metaproteomic findings regarding bacterial and human proteins, and its associations, we proposed a biliary tract carcinogenesis model. We are aware of that our proteomic analysis is a snapshot of established PDAC cases, and the propose bacteria-induced carcinogenesis model for the biliary tract is still speculative, not validated, and therefore must be interpreted with caution (**Figure 3**). The model initiates with unique or multiple dysbiotic factors that promote repeated inflammatory events (80), classically described as stones in the biliary tract, tobacco use, obesity, diabetes mellitus or genetic factors (81). Those promoting factors change the usual biliary tract bacteria composition expected for that individual, shaped according to diet, genetic background, sex, race, age, etc. Promoting factors can create many unusual microbiotas for that individual; though the specific carcinogenic dysbiotic microbiota has a reduced diversity as a sign of competition fostered by highly elevated synthesis of antibiotic products, and qualitative and quantitative differences in bacterial proteins associated with quorum sensing and biofilm formation, such as zeatin and surfactin. High levels of antibiotics maintain the dysbiotic environment, added to the antibiotic effect of surfactin, the latter also selecting, through inhibition, bacteria for biofilm formation. Bacterial species capable of biofilm formation will promote the synthesis of IL-8 by biliary tract epithelial cells. Fragments of IL-8 with antibiotic potential also contribute to maintaining dysbiosis, while the whole protein exerts its pro-neoplastic function of epithelial cellular survival, proliferation, angiogenesis, invasion and stem cell phenotype. The described scenario in conjunction with low levels or absence of zeatin, an anti-neoplastic protein, facilitates the progression of epithelial changes from low-grade dysplasia to adenocarcinoma, through mutation aggregation (82–84). The dysbiotic and harmful epithelial microenvironment needs to continue for a long but unspecified period of time to transform a benign epithelial cell into a malignant one, a period in which the molecular characteristic may be detected.

There are no clear indications for sample size calculations in proteomics research, and the results from a specific protein extraction method, imputation pipeline, and bioinformatic analysis must be validated in further *in vitro* and *in vivo* investigations. Microbiota analysis and dysbiosis alone have not answered the question of the bacteria-induced pathology model. Future research may concentrate on improving the throughput of protein identification from complex biological fluids like bile and consider a combined microbiota and metaproteomic approach to analyze bacterial communities and bacterial and human proteins. It is necessary to start thinking of a change in the dysbiosis paradigm, as we hypothesized dysbiosis is not a specific bacterial composition but rather a harmful protein microenvironment that can be created by several "dysbiotic" microbiotas.

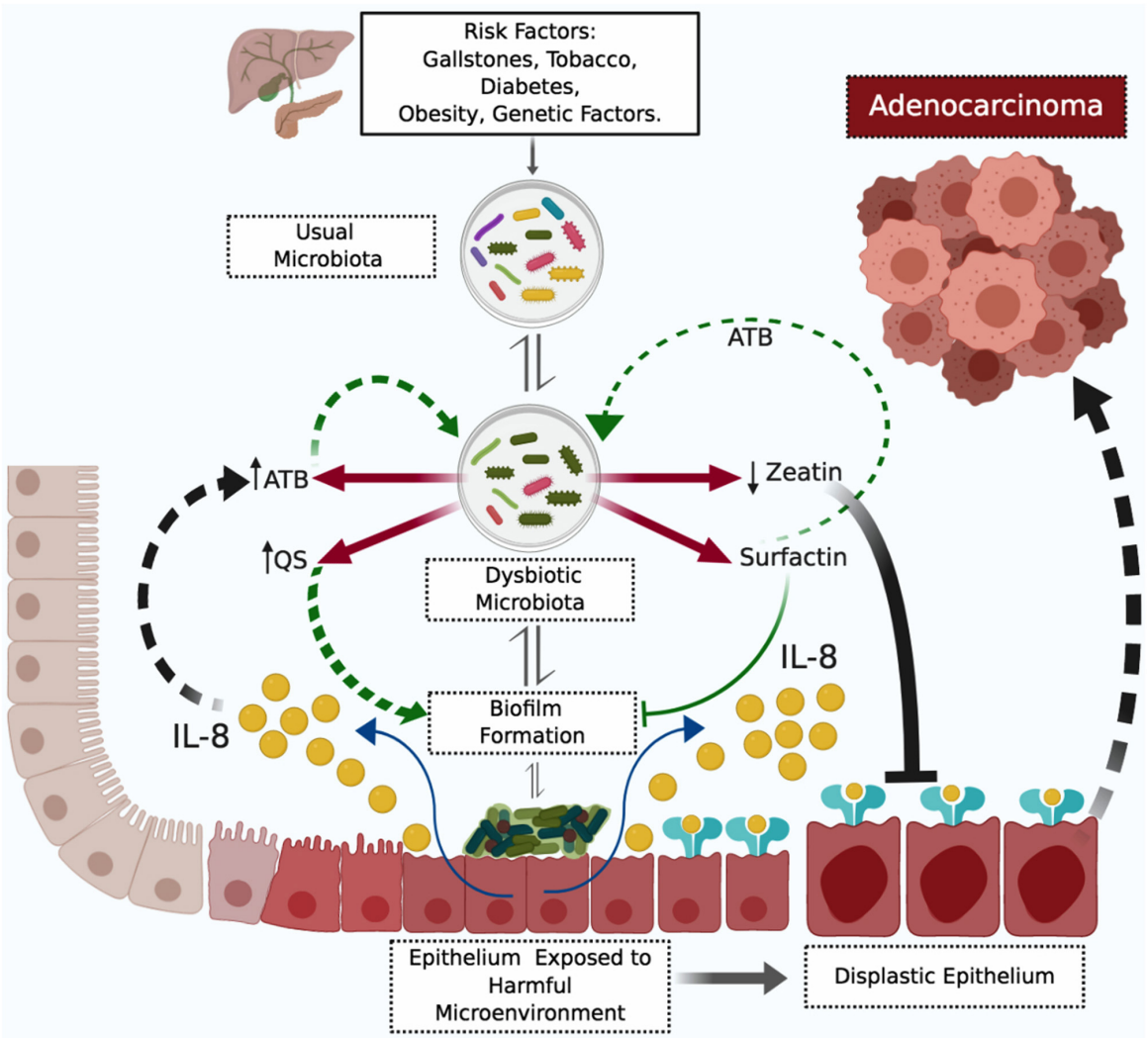


FIGURE 3 | Proposed Model of Biliary Tract Bacterial-Induced Carcinogenesis. The figure depicts a harmful microenvironment originating from prokaryote and eukaryote interaction. The harmful microenvironment initiates with a dysbiotic microbiota product of repetitive inflammatory processes induced by risk factors. The dysbiotic microbiota is specific for low levels of zeatin and high levels of antibiotics (ATB) and surfactin. Surfactin selectively inhibits bacterial biofilm formation—to date—without a molecular explanation for this selectivity. Bacterial biofilm formation stimulates IL-8 (interleukin 8) pro-neoplastic cytokine synthesis by biliary tract epithelial cells. Antibiotics, surfactin, and fragments of IL-8 with antibiotic properties perpetuate the dysbiotic microenvironment. Mutations accumulate in epithelial cells and IL-8 promotes the progression of dysplastic changes to adenocarcinoma, in low zeatin anti-neoplastic protein levels. ATB, antibiotics; IL-8, interleukin 8; QS, quorum sensing proteins.

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: ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020151.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committee—Universidad CES. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AA and NC-C design the study. DD, OP, and AA participated in sample acquisition. MS-J, NC-C, and AA participated in the laboratory procedures, protein analysis, and biological pathways interpretation. AA and NC-C wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Instituto Colombiano de Medicina Tropical and by CES University grant INV.022019.003.

REFERENCES

1. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology.* (2013) 144:1252–61. doi: 10.1053/j.gastro.2013.01.068
2. Ouaïssi M, Giger U, Louis G, Sielezneff I, Farges O, Sastre B. Ductal adenocarcinoma of the pancreatic head: a focus on current diagnostic and surgical concepts. *World J Gastroenterol.* (2012) 18:3058–69. doi: 10.3748/wjg.v18.i24.3058
3. Zhang Q, Zeng L, Chen Y, Lian G, Qian C, Chen S, et al. Pancreatic cancer epidemiology, detection, and management. *Gastroenterol Res Pract.* (2016) 2016:8962321. doi: 10.1155/2016/8962321
4. Zhang L, Sanagapalli S, Stoita A. Challenges in diagnosis of pancreatic cancer. *World J Gastroenterol.* (2018) 24:2047–60. doi: 10.3748/wjg.v24.i19.2047
5. Sener SF, Fremgen A, Menck HR, Winchester DP. Pancreatic cancer: a report of treatment and survival trends for 100,313 patients diagnosed from 1985-1995, using the national cancer database. *J Am Coll Surg.* (1999) 189:1–7. doi: 10.1016/S1072-7515(99)00075-7
6. Simard EP, Ward EM, Siegel R, Jemal A. Cancers with increasing incidence trends in the united states: 1999 through 2008. *CA Cancer J Clin.* (2012) 62:118–28. doi: 10.3322/caac.20141
7. Luo G, Zhang Y, Guo P, Ji H, Xiao Y, Li K. Global patterns and trends in pancreatic cancer incidence: age, period, and birth cohort analysis. *Pancreas.* (2019) 48:199–208. doi: 10.1097/MPA.0000000000001230
8. Vick AD, Hery DN, Markowiak SF, Brunicardi FC. Closing the disparity in pancreatic cancer outcomes: a Closer look at nonmodifiable factors and their potential use in treatment. *Pancreas.* (2019) 48:242–9. doi: 10.1097/MPA.0000000000001238
9. Camara SN, Yin T, Yang M, Li X, Gong Q, Zhou J, et al. High risk factors of pancreatic carcinoma. *Xuebao Yixue Yingdewen Ban.* (2016) 36:295–304.
10. Hollingworth R, Grand RJ. Modulation of DNA damage and repair pathways by human tumour viruses. *Viruses.* (2015) 7:2542–91. doi: 10.3390/v7052542
11. Vergara D, Simeone P, Damato M, Maffia M, Lanuti P, Trerotola M. The cancer microbiota: EMT and inflammation as shared molecular mechanisms associated with plasticity and progression. *J Oncol.* (2019) 2019:1253727. doi: 10.1155/2019/1253727
12. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. *Nat Rev Cancer.* (2013)13:759–71. doi: 10.1038/nrc3611
13. Avilés-Jiménez F, Guitron A, Segura-López F, Méndez-Tenorio A, Iwai S, Hernández-Guerrero A, et al. Microbiota studies in the bile duct strongly suggest a role for helicobacter pylori in extrahepatic cholangiocarcinoma. *Clin Microbiol Infect.* (2016) 22:178.e11-178.e22. doi: 10.1016/j.cmi.2015.10.008
14. Pereira P, Aho V, Arola J, Boyd S, Jokelainen K, Paulin L, et al. Bile microbiota in primary sclerosing cholangitis: impact on disease progression and development of biliary dysplasia. *PLoS ONE.* (2017) 12:e0182924. doi: 10.1371/journal.pone.0182924
15. Ye F, Shen H, Li Z, Meng F, Li L, Yang J, et al. Influence of the biliary system on biliary bacteria revealed by bacterial communities of the human biliary and upper digestive tracts. *PLoS ONE.* (2016) 11:e0150519. doi: 10.1371/journal.pone.0150519
16. Bridgewater JA, Goodman KA, Kalyan A, Mulcahy MF. Biliary tract cancer: epidemiology, radiotherapy, and molecular profiling. *Am Soc Clin Oncol Educ.* (2016) 35:e194-203. doi: 10.1200/EDBK_160831
17. Liou GY, Bastea L, Fleming A, Döppler H, Edenfield BH, Dawson DW, et al. The presence of interleukin-13 at pancreatic ADM/PanIN lesions alters macrophage populations and mediates pancreatic tumorigenesis. *Cell Rep.* (2017) 19:1322–33. doi: 10.1016/j.celrep.2017.04.052
18. Carrer A, Trefely S, Zhao S, Campbell SL, Norgard RJ, Schultz KC, et al. Acetyl-CoA metabolism supports multistep pancreatic tumorigenesis. *Cancer Discov.* (2019) 9:416–35. doi: 10.1158/2159-8290.CD-18-0567
19. Khan AS, Dageforde LA. Cholangiocarcinoma. *Surg Clin North Am.* (2019) 99:315–35. doi: 10.1016/j.suc.2018.12.004
20. Pushalkar S, Hundeyin M, Daley D, Zambirinis CP, Kurz E, Mishra A, et al. The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. *Cancer Discov.* (2018) 8:403–16. doi: 10.1158/2159-8290.CD-17-1134
21. Riquelme E, Zhang Y, Zhang L, Montiel M, Zoltan M, Dong W, et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell.* (2019) 178:795–806.e12. doi: 10.1016/j.cell.2019.07.008
22. Fan X, Alekseyenko AV, Wu J, Peters BA, Jacobs EJ, Gapstur SM, et al. Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. *Gut.* (2016) 67:120–7. doi: 10.1158/1538-7445.AM2016-4350
23. Farrell JJ, Zhang L, Zhou H, Chia D, Elashoff D, Akin D, et al. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. *Gut.* (2012) 61:582–8. doi: 10.1136/gutjnl-2011-300784
24. Pan S, Brentnall TA, Chen R. Proteomics analysis of bodily fluids in pancreatic cancer. *Proteomics.* (2015) 15:2705–15. doi: 10.1002/pmic.201400476
25. Chen KT, Kim PD, Jones KA, Devarajan K, Patel BB, Hoffman JP, et al. Potential prognostic biomarkers of pancreatic cancer. *Pancreas.* (2014) 43:22–7. doi: 10.1097/MPA.0b013e3182a6867e
26. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* (2019) 47:W191–8. doi: 10.1093/nar/gkz369
27. Kanehisa M, Sato Y. KEGG mapper for inferring cellular functions from protein sequences. *Protein Sci Publ Protein Soc.* (2019) 29:28–35. doi: 10.1002/pro.3711
28. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in kEGG. *Nucleic Acids Res.* (2014) 42(Database issue):D199–205. doi: 10.1093/nar/gkt1076
29. Tan AA, Azman SN, Abdul Rani NR, Kua BC, Sasidharan S, Kiew LV, et al. Optimal protein extraction methods from diverse sample types for protein profiling by using two-Dimensional electrophoresis (2DE). *Trop Biomed.* (2011) 28:620–9.
30. Wieczorek S, Combes F, Lazar C, Giai Gianetto Q, Gatto L, Dorffer A, et al. DAPAR & proStaR: software to perform statistical analyses in quantitative discovery proteomics. *Bioinforma Oxf Engl.* (2017) 33:135–6. doi: 10.1093/bioinformatics/btw580
31. Kanehisa M, Sato Y, Morishima K. Blastkoala and Ghostkoala: Kegg tools for functional characterization of genome and metagenome sequences. *J Mol Biol.* (2016) 428:726–31. doi: 10.1016/j.jmb.2015.11.006
32. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped uIs and aPIs. *Nucleic Acids Res.* (2019) 47:W199–205. doi: 10.1093/nar/gkz401
33. Oladele Ogunseitan. Cross-species interactions among prokaryotes. In: *Microbial Diversity Form and Function in Prokaryotes*. Oxford, UK: Blackwell Science Ltd (2005).
34. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol.* (2020) 38:685–88. doi: 10.1038/s41587-020-0548-6
35. Wilmes P, Bond PL. Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol.* (2006) 14:92–7. doi: 10.1016/j.tim.2005.12.006
36. Blakeley-Ruiz JA, Erickson AR, Cantarel BL, Xiong W, Adams R, Jansson JK, et al. Metaproteomics reveals persistent and phylum-redundant metabolic functional stability in adult human gut microbiomes of crohn's remission patients despite temporal variations in microbial taxa, genomes, and proteomes. *Microbiome.* (2019) 7:18. doi: 10.1186/s40168-019-0631-8
37. Peters DL, Wang W, Zhang X, Ning Z, Mayne J, Figeys D. Metaproteomic and metabolomic approaches for characterizing the gut microbiome. *Proteomics.* (2019) 19:e1800363. doi: 10.1002/pmic.201800363
38. Patel K, Singh M, Gowda H. Bioinformatics methods to deduce biological interpretation from proteomics data. *Methods Mol Biol Clifton NJ.* (2017) 1549:147–61. doi: 10.1007/978-1-4939-6740-7_12
39. Mukaida N, Harada A, Yasumoto K, Matsushima K. Properties of pro-inflammatory cell type-specific leukocyte chemotactic cytokines, interleukin 8 (IL-8) and monocyte chemotactic and activating factor (MCAF). *Microbiol Immunol.* (1992) 36:773–89. doi: 10.1111/j.1348-0421.1992.tb02080.x
40. Russo RC, Garcia CC, Teixeira MM, Amaral FA. The cXCL8/IL-8 chemokine family and its receptors in inflammatory diseases. *Expert Rev Clin Immunol.* (2014) 10:593–619. doi: 10.1586/1744666X.2014.894886

41. Young RS, Wiles BM, McGee DW. IL-22 enhances tNF-α- and iL-1-Induced cXCL8 responses by intestinal epithelial cell lines. *Inflammation.* (2017) 40:1726–34. doi: 10.1007/s10753-017-0614-5
42. Kasahara T, Mukaida N, Yamashita K, Yagisawa H, Akahoshi T, Matsushima K. IL-1 and tNF-alpha induction of iL-8 and monocyte chemotactic and activating factor (MCAF) mRNA expression in a human astrocytoma cell line. *Immunology.* (1991) 74:60–7.
43. Waugh DJJ, Wilson C. The interleukin-8 pathway in cancer. *Clin Cancer Res Off J Am Assoc Cancer Res.* (2008) 14:6735–41. doi: 10.1158/1078-0432.CCR-07-4843
44. Alfaro C, Sanmamed MF, Rodríguez-Ruiz ME, Teijeira Á, Oñate C, González Á, et al. Interleukin-8 in cancer pathogenesis, treatment and follow-up. *Cancer Treat Rev.* (2017) 60:24–31. doi: 10.1016/j.ctrv.2017.08.004
45. Liu Q, Li A, Tian Y, Wu JD, Liu Y, Li T, et al. The cXCL8-CXCR1/2 pathways in cancer. *Cytokine Growth Factor Rev.* (2016) 31:61–71. doi: 10.1016/j.cytogfr.2016.08.002
46. Kamohara H, Takahashi M, Ishiko T, Ogawa M, Baba H. Induction of interleukin-8 (CXCL-8) by tumor necrosis factor-alpha and leukemia inhibitory factor in pancreatic carcinoma cells: impact of cXCL-8 as an autocrine growth factor. *Int J Oncol.* (2007) 31:627–32. doi: 10.3892/ijo.31.3.627
47. Ebrahimi B, Tucker SL, Li D, Abbruzzese JL, Kurzrock R. Cytokines in pancreatic carcinoma: correlation with phenotypic characteristics and prognosis. *Cancer.* (2004) 101:2727–36. doi: 10.1002/cncr.20672
48. Sun Q, Li F, Sun F, Niu J. Interleukin-8 is a prognostic indicator in human hilar cholangiocarcinoma. *Int J Clin Exp Pathol.* (2015) 8:8376–84.
49. Chen L, Fan J, Chen H, Meng Z, Chen Z, Wang P, et al. The iL-8/CXCR1 axis is associated with cancer stem cell-like properties and correlates with clinical prognosis in human pancreatic cancer cases. *Sci Rep.* (2014) 4:5911. doi: 10.1038/srep05911
50. Fang Y, Saiyin H, Zhao X, Wu Y, Han X, Lou W. IL-8-Positive tumor-Infiltrating inflammatory cells are a novel prognostic marker in pancreatic ductal adenocarcinoma patients. *Pancreas.* (2016) 45:671–8. doi: 10.1097/MPA.0000000000000520
51. Peyyala R, Kirakodu SS, Novak KF, Ebersole JL. Oral microbial biofilm stimulation of epithelial cell responses. *Cytokine.* (2012) 58:65–72. doi: 10.1016/j.cyto.2011.12.016
52. Ramage G, Lappin DF, Millhouse E, Malcolm J, Jose A, Yang J, et al. The epithelial cell response to health and disease associated oral biofilm models. *J Periodontal Res.* (2017) 52:325–33. doi: 10.1111/jre.12395
53. Ebersole JL, Peyyala R, Gonzalez OA. Biofilm-induced profiles of immune response gene expression by oral epithelial cells. *Mol Oral Microbiol.* (2019) 34:1. doi: 10.1111/omi.12251
54. Björstad A, Fu H, Karlsson A, Dahlgren C, Bylund J. Interleukin-8-derived peptide has antibacterial activity. *Antimicrob Agents Chemother.* (2005) 49:3889–95. doi: 10.1128/AAC.49.9.3889-3895.2005
55. Guo Q, Jian Z, Jia B, Chang L. CXCL7 promotes proliferation and invasion of cholangiocarcinoma cells. *Oncol Rep.* (2017) 37:1114–22. doi: 10.3892/or.2016.5312
56. Dufies M, Giuliano S, Viotti J, Borchiellini D, Cooley LS, Ambrosetti D, et al. CXCL7 is a predictive marker of sunitinib efficacy in clear cell renal cell carcinomas. *Br J Cancer.* (2017) 117:947–53. doi: 10.1038/bjc.2017.276
57. Desurmont T, Skrypek N, Duhamel A, Jonckheere N, Millet G, Leteurtre E, et al. Overexpression of chemokine receptor cXCR2 and ligand cXCL7 in liver metastases from colon cancer is correlated to shorter disease-free and overall survival. *Cancer Sci.* (2015) 106:262–9. doi: 10.1111/cas.12603
58. Li Y, Yu HP, Zhang P. CCL15 overexpression predicts poor prognosis for hepatocellular carcinoma. *Hepatol Int.* (2016) 10:488–92. doi: 10.1007/s12072-015-9683-4
59. Li Y, Wu J, Zhang P. CCL15/CCR1 axis is involved in hepatocellular carcinoma cells migration and invasion. *Tumour Biol J Int Soc Oncodevelopmental Biol Med.* (2016) 37:4501–7. doi: 10.1007/s13277-015-4287-0
60. Yang S-M, Li S-Y, Hao-Bin Y, Lin-Yan X, Sheng X. IL-11 activated by lnc-ATB promotes cell proliferation and invasion in esophageal squamous cell cancer. *Biomed Pharmacother Biomedecine Pharmacother.* (2019) 114:108835. doi: 10.1016/j.biopha.2019.108835
61. Putoczki TL, Thiem S, Loving A, Busuttil RA, Wilson NJ, Ziegler PK, et al. Interleukin-11 is the dominant iL-6 family cytokine during gastrointestinal tumorigenesis and can be targeted therapeutically. *Cancer Cell.* (2013) 24:257–71. doi: 10.1016/j.ccr.2013.06.017
62. Kamada-Nobusada T, Sakakibara H. Molecular basis for cytokinin biosynthesis. *Phytochemistry.* (2009) 70:444–9. doi: 10.1016/j.phytochem.2009.02.007
63. Akiyoshi DE, Regier DA, Gordon MP. Cytokinin production by agrobacterium and pseudomonas spp. *J Bacteriol.* (1987) 169:4242–8. doi: 10.1128/JB.169.9.4242-4248.1987
64. Kakimoto T. Biosynthesis of cytokinins. *J Plant Res.* (2003) 116:233–9. doi: 10.1007/s10265-003-0095-5
65. Voller J, Zatloukal M, Lenobel R, Dolezal K, Béres T, Krystof V, et al. Anticancer activity of natural cytokinins: a structure-activity relationship study. *Phytochemistry.* (2010) 71:1350–9. doi: 10.1016/j.phytochem.2010.04.018
66. Siwak E, Jewginski M, Kustrzeba-Wójcicka I. Biological activity of surfactins - a case of a biosurfactant produced by *Bacillus subtilis* pCM (1949). *Acta Biochim Pol.* (2015) 62:875–8. doi: 10.18388/abp.2015_1149
67. Das P, Mukherjee S, Sen R. Antimicrobial potential of a lipopeptide biosurfactant derived from a marine bacillus circulans. *J Appl Microbiol.* (2008) 104:1675–84. doi: 10.1111/j.1365-2672.2007.03701.x
68. Loiseau C, Schlusselhuber M, Bigot R, Bertaux J, Berjeaud J-M, Verdon J. Surfactin from *Bacillus subtilis* displays an unexpected anti-Legionella activity. *Appl Microbiol Biotechnol.* (2015) 99:5083–93. doi: 10.1007/s00253-014-6317-z
69. Zhao H, Shao D, Jiang C, Shi J, Li Q, Huang Q, et al. Biological activity of lipopeptides from bacillus. *Appl Microbiol Biotechnol.* (2017) 101:5951–60. doi: 10.1007/s00253-017-8396-0
70. Yuan L, Zhang S, Wang Y, Li Y, Wang X, Yang Q. Surfactin inhibits membrane fusion during invasion of epithelial cells by enveloped viruses. *J Virol.* (2018) 92:18. doi: 10.1128/JVI.00809-18
71. Liu J, Li W, Zhu X, Zhao H, Lu Y, Zhang C, et al. Surfactin effectively inhibits staphylococcus aureus adhesion and biofilm formation on surfaces. *Appl Microbiol Biotechnol.* (2019) 103:4565–74. doi: 10.1007/s00253-019-09808-w
72. Bais HP, Fall R, Vivanco JM. Biocontrol of bacillus subtilis against infection of arabidopsis roots by pseudomonas syringae is facilitated by biofilm formation and surfactin production. *Plant Physiol.* (2004) 134:307–19. doi: 10.1104/pp.103.028712
73. Moryl M, Spetana M, Dziubek K, Paraszkiewicz K, Rózalska S, Płaza GA, et al. Antimicrobial, antiadhesive and antibiofilm potential of lipopeptides synthesised by bacillus subtilis, on uropathogenic bacteria. *Acta Biochim Pol.* (2015) 62:725–32. doi: 10.18388/abp.2015_1120
74. Netzker T, Flak M, Krespach MK, Stroe MC, Weber J, Schroeckh V, et al. Microbial interactions trigger the production of antibiotics. *Curr Opin Microbiol.* (2018) 45:117–23. doi: 10.1016/j.mib.2018.04.002
75. Abrudan MI, Smakman F, Grimbergen AJ, Westhoff S, Miller EL, van Wezel GP, et al. Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc Natl Acad Sci USA.* (2015) 112:11054–9. doi: 10.1073/pnas.1504076112
76. López D, Vlamakis H, Kolter R. Biofilms. *Cold Spring Harb Perspect Biol.* (2010) 2:a000398. doi: 10.1101/cshperspect.a000398
77. Sakuragi Y, Kolter R. Quorum-sensing regulation of the biofilm matrix genes (pel) of pseudomonas aeruginosa. *J Bacteriol.* (2007) 189:5383–6. doi: 10.1128/JB.00137-07
78. Balkwill F, Mantovani A. Inflammation and cancer: back to virchow? *Lancet Lond Engl.* (2001) 357:539–45. doi: 10.1016/S0140-6736(00)04046-0
79. Francescone R, Hou V, Grivennikov SI. Microbiome, inflammation, and cancer. *Cancer J Sudbury Mass.* (2014) 20:181–9. doi: 10.1097/PPO.0000000000000048
80. Shadhu K, Xi C. Inflammation and pancreatic cancer: an updated review. *Saudi J Gastroenterol Off J Saudi Gastroenterol Assoc.* (2019) 25:3–13. doi: 10.4103/sjg.SJG_390_18

81. Kleeff J, Korc M, Apte M, La Vecchia C, Johnson CD, Biankin AV, et al. Pancreatic cancer. *Nat Rev Dis Primer.* (2016) 2:16022. doi: 10.1038/nrdp.2016.22
82. Barreto SG, Dutt A, Chaudhary A. A genetic model for gallbladder carcinogenesis and its dissemination. *Ann Oncol Off J Eur Soc Med Oncol ESMO.* (2014) 25:1086–97. doi: 10.1093/annonc/mdu006
83. Macgregor-Das AM, Iacobuzio-Donahue CA. Molecular pathways in pancreatic carcinogenesis. *J Surg Oncol.* (2013) 107:8–14. doi: 10.1002/jso.23213
84. Makohon-Moore A, Iacobuzio-Donahue CA. Pancreatic cancer biology and genetics from an evolutionary perspective. *Nat Rev Cancer.* (2016) 16:553–65. doi: 10.1038/nrc.2016.66

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 © 2020 Arteta, Sánchez-Jiménez, Dávila, Palacios and Cardona-Castro. 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 Oncology

ORIGINAL RESEARCH
published: 25 August 2020
doi: 10.3389/fonc.2020.01601

Schwann Cell Stimulation of Pancreatic Cancer Cells: A Proteomic Analysis

*Aysha Ferdoushi[1,2,3], Xiang Li[1,2], Nathan Griffin[1,2], Sam Faulkner[1,2], M. Fairuz B. Jamaluddin[1,2], Fangfang Gao[1,2], Chen Chen Jiang[2,4], Dirk F. van Helden[1,2], Pradeep S. Tanwar[1,2], Phillip Jobling[1,2] and Hubert Hondermarck[1,2]**

[1] School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, University of Newcastle, Callaghan, NSW, Australia, [2] Hunter Medical Research Institute, University of Newcastle, New Lambton, NSW, Australia, [3] Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh, [4] School of Medicine and Public Health, The University of Newcastle, Callaghan, NSW, Australia

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Helmut Friess,
Technical University of Munich, Germany
Dileep Kumar,
The University of Utah, United States
Marco Artico,
Sapienza University of Rome, Italy

***Correspondence:**
Hubert Hondermarck
hubert.hondermarck@newcastle.edu.au

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 19 April 2020
Accepted: 23 July 2020
Published: 25 August 2020

Citation:
Ferdoushi A, Li X, Griffin N, Faulkner S, Jamaluddin MFB, Gao F, Jiang CC, van Helden DF, Tanwar PS, Jobling P and Hondermarck H (2020) Schwann Cell Stimulation of Pancreatic Cancer Cells: A Proteomic Analysis. Front. Oncol. 10:1601. doi: 10.3389/fonc.2020.01601

Schwann cells (SCs), the glial component of peripheral nerves, have been identified as promoters of pancreatic cancer (PC) progression, but the molecular mechanisms are unclear. In the present study, we aimed to identify proteins released by SCs that could stimulate PC growth and invasion. Proteomic analysis of human primary SC secretome was performed using liquid chromatography–tandem mass spectrometry, and a total of 13,796 unique peptides corresponding to 1,470 individual proteins were identified. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment were conducted using the Database for Annotation, Visualization, and Integrated Discovery. Metabolic and cell–cell adhesion pathways showed the highest levels of enrichment, a finding in line with the supportive role of SCs in peripheral nerves. We identified seven SC-secreted proteins that were validated by western blot. The involvement of these SC-secreted proteins was further demonstrated by using blocking antibodies. PC cell proliferation and invasion induced by SC-conditioned media were decreased using blocking antibodies against the matrix metalloproteinase-2, cathepsin D, plasminogen activator inhibitor-1, and galectin-1. Blocking antibodies against the proteoglycan biglycan, galectin-3 binding protein, and tissue inhibitor of metalloproteinases-2 decreased only the proliferation but not the invasion of PC cells. Together, this study delineates the secretome of human SCs and identifies proteins that can stimulate PC cell growth and invasion and therefore constitute potential therapeutic targets.

Keywords: pancreatic cancer, Schwann cells, secretome, cancer cell proliferation and invasion, LC-MS/MS, therapeutic targets

Abbreviations: cat. no., catalog number; CM, conditioned media; DAVID, Database for Annotation, Visualization, and Integrated Discovery; ECM, extracellular matrix; GAD, Genetic Association Database; Gal-1, galectin-1; Gal-3BP, galectin-3 binding protein; GO, Gene Ontology; IGFBP, insulin-like growth factor-binding proteins; IHC, Immunohistochemistry; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS/MS, liquid chromatography–tandem mass spectrometry; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; PC, pancreatic cancer; RFU, relative fluorescence unit; SC, Schwann cell; SCM, Schwann cell medium; SD, standard deviation; SF, serum-free; TCGA, The Cancer Genome Atlas; TIMP-2, tissue inhibitor of metalloproteinases-2; TMA, tissue micro array; WB, western blot.

INTRODUCTION

Pancreatic cancer (PC) is one of the most lethal malignancies (1) and is predicted to become the second leading cause of cancer-related death by 2030 (2). Eighty-five percent of PC cases are ductal adenocarcinomas with a 5-year survival rate less than 7% (3). The poor patient survival is attributed to late-stage diagnosis, high incidence of local recurrence, development of distant metastases, and therapeutic resistance (4, 5). In addition, there is currently no targeted therapy for PC, and therefore the identification of potential therapeutic targets is essential.

Schwann cells (SCs) are the major glial component in the peripheral nervous system (6). SCs maintain neuronal homeostasis through the regulation of cell growth, survival, and repair (7, 8). The primordial role of SCs is myelination (9), however SCs have recently been implicated in several malignancies including pancreatic (10–12), prostate (13), lung (14), oral (15), and cervical (16) cancers. In PC, SCs are involved in the initiation of disease, and their presence is associated with increased perineural invasion, the process by which cancer cells invade nerves (10). SCs guide cancer cells toward nerves via the production of neural cell adhesion molecule 1 (NCAM 1) that promotes perineural invasion (11). SCs have also been shown to initiate epithelial–mesenchymal transition and support metastatic spread (12), and SCs have been reported to mask cancer-related pain, resulting in a prolonged asymptomatic phase and delayed cancer diagnosis (17). Additionally, SC-derived interleukin 6 has been reported to augment PC cell migration and invasion (18). Although few SC-secreted cytokines (19) and adhesion molecules (20) have been described, the secretome of SC and its impact in PC remain largely unknown.

In the present study, we have profiled the secretome of human SCs using liquid chromatography–tandem mass spectrometry (LC-MS/MS) and investigated the role of several identified proteins in the stimulation of PC growth and invasion. These secreted proteins may constitute new therapeutic targets for PC.

MATERIALS AND METHODS

Human SC Culture

Primary human SCs, obtained from the spinal nerve cells of a healthy donor, were purchased from ScienCell (cat. no. 1700, CA, United States) and maintained (maximum of 10 passages) according to manufacturer instructions described previously (21). Briefly, T-75 culture flasks were coated with 10 mg/mL poly-L-lysine (cat. no., 0413, ScienCell) and incubated overnight at 37°C. Cells were seeded at 5,000 cells/cm^2 on the poly-L-lysine–coated flask after washing the vessel twice with sterile milli-Q water. Cells were grown in complete SC medium (SCM, cat. no., 1701, ScienCell), supplemented with 5% fetal bovine serum (FBS, cat. no., 0025, ScienCell), 1% SC growth supplement cocktail (SCGS, cat. no., 1752, ScienCell), and 1% penicillin/streptomycin (P/S, cat. no., 0503, ScienCell) in a humidified incubator at 37°C with 5% CO_2. SCs were characterized by immunoblotting using antibodies against human SC marker proteins, SOX10, and p75 (**Supplementary Figure S1**).

Pancreatic ductal adenocarcinoma cells, PANC-1 and MIA PaCa-2, were obtained from the American Type Culture Collection (ATCC, Manassas, VA, United States) and maintained in Dulbecco modified eagle medium (cat. no., ATC302002, ATCC) supplemented with 10% (vol/vol) FBS (JRH Biosciences, St. Louis, MO, United States) and 2 mM L-glutamine in a humidified incubator at 37°C with 5% (vol/vol) CO_2.

SC Conditioned Media Preparation

SCs were grown to 70–80% confluency in SCM and washed three times with sterile phosphate-buffered saline (PBS) (Invitrogen, CA, United States) and once with serum-free (SF) media. SCs were then incubated in SF media for 20 h, after which SC-conditioned media (SC-CM) was collected, centrifuged (1,000 × *g* at 4°C for 10 min), and the supernatant was filtered through a 0.22-μm nylon filter (Merck Millipore, MA, United States) to remove any cell debris or floating cells. SC-CM was further centrifuged (4,000 × *g* at 4°C for 30 min) to concentrate using a 3-kDa cutoff Amicon Ultra-15 filter unit (Merck Millipore) until the media was concentrated 400-fold. The recovered SC-CM concentrate was stored at −80°C. An outline of SC-CM collection and concentration workflow are shown in **Figure 1A**.

Mass Spectrometry–Based Proteomic Analysis of SC Secretome

Discovery proteomic analysis was performed by LC-MS/MS to describe the secretome profile of primary human SCs. Experimental protocol outlined in **Figure 1B** has been previously described (21) with additional sample preparation steps added for the secretome analysis. The secretome of SCs have been analyzed in an unstimulated state, i.e., without previous exposure to PC cells.

Sample Preparation for LC-MS/MS

Two hundred micrograms of secreted proteins was measured by bicinchoninic acid (BCA) assay (Pierce, Thermo Fisher Scientific, IL, United States) and dissolved in urea (6 M urea 2 M thiourea) buffer followed by reduction step using 10 mM dithiothreitol (30 min at room temperature). The samples were subsequently alkylated using 20 mM iodoacetamide (30 min at room temperature in the dark). Proteins were digested using 1:40 ratio Lys-C/trypsin (cat. no., VA1170, Promega, Madison, WI, United States) to protein concentration (3 h, room temperature, in the dark). The concentration of the urea was brought down to less than 1 M by adding 20 mM triethylammonium bicarbonate (pH 7.8) and incubated overnight at room temperature. Peptides were desalted and cleaned up using a Visiprep^{TM} vacuum manifold (12-port, cat. no. 57030-U, Sigma-Aldrich, St. Louis, MO, United States) coupled with Empore C18 solid-phase extraction cartridge (4 mm/1 mL) according to manufacturer instructions.

LC-MS/MS

A Dionex UltiMate 3000 nanoLC system (Thermo Fisher Scientific) with a 15-cm EASY-Spray Column was used to

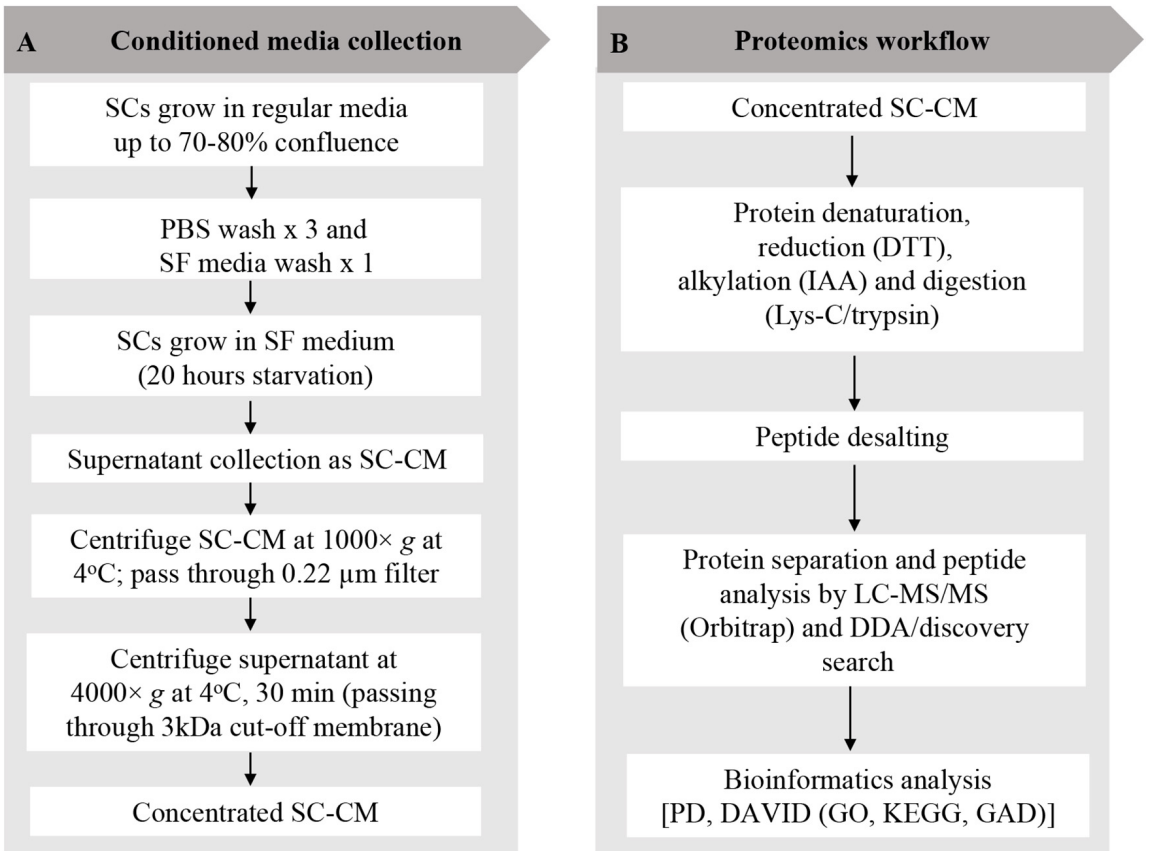


FIGURE 1 | Schwann cell–conditioned media (SC-CM) collection and proteomic workflow. **(A)** For SC-CM collection, SCs were grown to 70–80% confluence. Cells were washed three times with sterile PBS and once with SF media before incubation in SF media for additional 20 h. SC-CM were then collected and centrifuged (1,000 × *g*, 4°C, 10 min), and the supernatant was filtered through a 0.22-µm nylon filter to remove any cell debris or floating cells. Collected supernatant was concentrated using 3-kDa cutoff Amicon Ultra-15 filter unit (4,000 × *g*, 4°C, 30 min). **(B)** Proteomic analysis of the secretome was performed using LC-MS/MS. Concentrated SC-CM was subjected to reduction (DTT), alkylation (IAA), and Lys-C/trypsin digestion before analysis in LC-MS/MS. DDA discovery search was performed to identify the total number of peptides and proteins. To profile the SC-secretome, functional clustering of the secreted proteins was performed using DAVID bioinformatics software and displayed in GO and KEGG. Identified proteins were classified based on disease classes using DAVID-GAD analysis. SC, Schwann cell; PBS, phosphate-buffered saline; SF, serum-free; SC-CM, Schwann cell–conditioned media; DTT, dithiothreitol; IAA, iodoacetamide; LC-MS/MS, liquid chromatography–tandem mass spectrometry; DDA, data-dependent acquisition; PD, Proteome Discoverer; DAVID, Database for Annotation, Visualization, and Integrated Discovery; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GAD, Genetic Association Database.

separate 500 ng of peptides using a 150-min gradient at a flow rate of 300 nL/min. Peptides were analyzed on Q-Exactive Plus Orbitrap (Thermo Fisher Scientific) mass spectrometer. Precursor scan of intact peptides was measured in the Orbitrap by scanning from m/z 400–2,000, with a resolution of 70 K with maximum ion injection time of 50 ms and automated gain control (AGC) target of 1E6. The 15 most intense multiply charged precursors were selected for HCD fragmentation with a normalized collision energy of 27.0 and then measured in the Orbitrap at a resolution of 35 K with maximum ion injection time of 120 ms, and AGC target was set at 2E5. Dynamic exclusion was set for 30 s.

MS Data Analysis

LC-MS/MS data were analyzed using Proteome Discoverer software v.2.1 (Thermo Fisher Scientific) and searched against the Uniprot human protein database (downloaded March 5, 2018, with a total of 71,773 entries). Precursor mass tolerance was set to 10 ppm, and fragment ion tolerance was 0.02 Da. Trypsin was designated as the digestion enzyme with two missed cleavages permitted. Carbamidomethylation on cysteine (+57.021 Da) was set as static modifications, and oxidation on methionine (+15.995 Da) was set as variable modification. Only high confidence identification, represented by at least two unique peptides, was included in the analysis. Protein confidence indicators were set at 1% false discovery rate criteria using a percolator. The precursor ion (MS) spectra were also manually validated using Xcalibur Software version 4.0.27.13 (Thermo Fisher Scientific) to avoid false-positive detection.

Proliferation Assay

Pancreatic cancer cells (PANC-1 and MIA PaCa-2) were starved for 24 h and seeded (∼5,000 cells/well) in a 96-well culture plate.

Cells were cultured in 200 μL SF media, serum-supplemented media (noted as FBS), and SC-CM, with or without specific blocking antibodies (6 μg/mL) (detailed information of the antibodies is listed in **Table 1**). Antibodies were used to block specific proteins in SC-CM. Of note, the commercial producers have reported blocking activity for the antibodies that we have used, but we have not independently confirmed it. Cells were cultured for 72 h and subsequently incubated with CellTiter-Blue® (Promega Corporation, cat. no., G8081) at 37°C for 4 h before recording fluorescence at 560/590 using a FLUOstar OPTIMA fluorescence plate reader (BMG Labtech, Durham, NC, Untied States). All the experiments were repeated three times.

Invasion Assay

Cell invasion assays were performed on serum-starved PC cells (PANC-1 and MIA PaCa-2) using the QCM ECM Cell Invasion Assay kit (cat. no., ECM554; Merck Millipore). The supplied 24-well assay plate contains upper invasion chamber inserts with 8-mm pore size membranes. The extracellular matrix (ECM) layer was rehydrated with 300 μL of prewarmed SF media for 30 min at room temperature. Serum-starved cells (∼60,000) were loaded into the Transwell chamber insert in 250 μL of SF media or SC-CM with or without specific antibody (6 μg/mL). Five hundred microliters of SF medium, serum-supplemented media (noted as FBS), and SC-CM, with or without specific antibody (6 μg/mL), was added to the lower chamber. After 24 h, invading cells were dislodged, and the fluorescence was recorded at 480/520 nm using a FLUOstar OPTIMA fluorescence plate reader (BMG Labtech) as described previously (22). All the experiments were repeated three times.

TABLE 1 | Antibodies used for WB.

No.	Antibody name	Source (working dilution)	Company (cat. no.)
Antibodies used to detect protein in SC-CM			
1	Gal-3BP	Goat polyclonal (1:700)	R&D System (AF2226)
2	MMP-2	Goat polyclonal (1:500)	R&D System (AF902)
3	Cathepsin D	Goat polyclonal (1:100)	R&D System (AF1014)
4	PAI-1	Goat polyclonal (1:100)	R&D System (AF1786)
5	Biglycan	Goat polyclonal (1:1,000)	R&D System (AF2667)
6	TIMP-2	Goat polyclonal (1:500)	R&D System (AF971)
7	Gal-1	Goat polyclonal (1:1,000)	R&D System (AF1152)
Antibodies used to detect protein in cell lysate			
8	Sox10	Mouse monoclonal (1:400)	Abcam (ab212843)
9	p75	Rabbit polyclonal (1:200)	Santa Cruz, (H-137): sc-8317
Antibodies used to detect loading control			
10	β-actin	Mouse monoclonal (1:2,000)	Abcam (ab8226)
11	β-actin	Rabbit polyclonal (1:2,000)	Abcam (ab8227)

Antibodies used to detect proteins both in SC-CM and cell lysate are presented here with source, dilution, company name, and cat. no. SC-CM, Schwann cell–conditioned media; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1; cat. no., catalog number.

Western Blotting

Concentrated SC-CM and cellular lysates underwent western blot (WB) analysis to detect specific proteins of interest from LC-MS/MS in the SC secretome and SCs, respectively. After collecting the SC-CM for secretome profiling and functional analysis, the remaining cells were washed three times with PBS, trypsinized, and collected by gently scrapping into PBS media. Cells were subsequently pelleted by centrifugation at 800 × *g* for 5 min. Total proteins were extracted from cell pellets using RIPA buffer [25 mM Tris–HCl (pH 7.6)], 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS)] (Thermo Fisher Scientific) and commercial protease inhibitor and phosphatase inhibitor cocktail tablets (Roche, Mannheim, Germany), aliquoted, and stored at −20°C. The total protein concentration of cell extracts and concentrated SC-CM was determined using a BCA assay (Pierce), according to the manufacturer's instructions.

Thirty micrograms of protein from each sample was resuspended in an equal volume of Laemmli buffer (Bio-Rad, Hercules, CA, United States). The cell extract or concentrated SC-CM was subjected to SDS–polyacrylamide gel electrophoresis under reducing conditions, and the separated proteins were transferred to 0.4-mm pore nitrocellulose membranes (Amersham, GE Healthcare Life Sciences, Pittsburgh, PA, United States). Blots were blocked with blocking buffer (LI-COR Biosciences, Lincoln, NE, United States) for 1 h at room temperature and then probed with antibodies against specific proteins (**Table 1**). Identical antibodies were used for both WB and functional analysis. β-Actin protein expression was used as loading control. All antibodies were diluted in blocking buffer (LI-COR Biosciences). After washing with PBS containing 0.1% Tween-20, membranes were probed with goat anti–mouse or goat anti–rabbit IR-Dye 670 or 800 cw labeled secondary antisera, and then washes were repeated after labeling. WB was imaged using the LI-COR Odyssey infrared imaging system (LI-COR Biosciences).

Pancreatic Tissue Samples and Immunohistochemistry

High-density tumor micro arrays (TMAs) were obtained from US Biomax Inc. (Maryland, MD, United States). The TMAs used (HPan-Ade170Sur-01) included a total of 99 pancreatic adenocarcinomas and 71 normal adjacent pancreatic tissues. For each specimen collected, informed consent was obtained from both the hospital and the individual. Discrete legal consent was obtained, and the rights to hold research uses for any purpose or further commercialized uses were waived. The study was approved by the University of Newcastle's Human Research Ethics Committee.

Immunohistochemistry (IHC) was performed as described previously (23). Following deparaffinization and rehydration of the TMA slides using standard procedures, heat-induced epitope retrieval was carried out in a low-pH, citrate-based antigen unmasking solution (catalog number H-3300, Vector Laboratories, California, CA, United States) by a decloaking chamber (Biocare, West Midlands, United Kingdom) at 95°C for

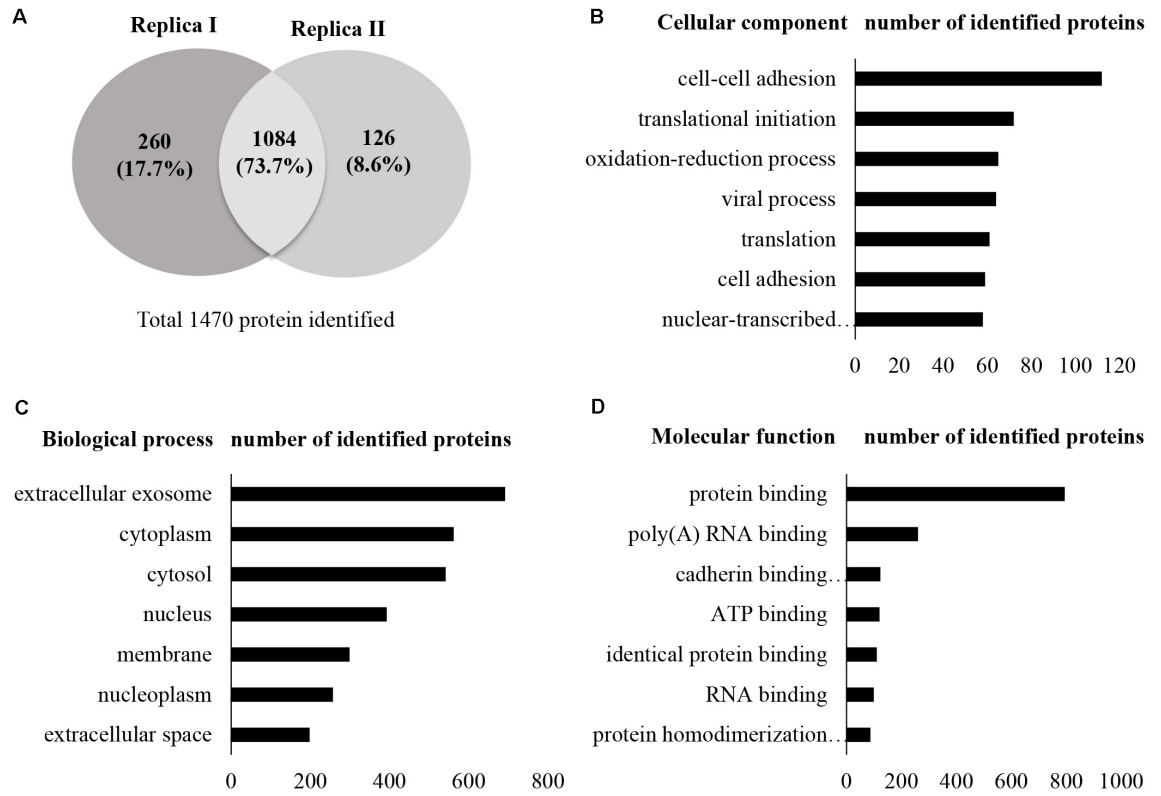


FIGURE 2 | Secretome profile and functional clustering of human SCs. **(A)** Venn diagram summarizing the overlap in protein identification between two independent proteomic analyzes of SC-CM, which identified 1,344 (replica I) and 1,210 (replica II) proteins, respectively. Additional details are listed in **Supplementary Tables S2–S4**. **(B–D)** DAVID bioinformatics resource was used to analyze the secretome of SCs in order to identify enriched biological terms in the common secreted proteins extracted from the MS analysis. **(B)** GO enrichment analysis of cellular components. Proteins localized in extracellular exosomes were highly enriched. **(C)** GO enrichment analysis of biological processes. Highly enriched SC secreted proteins are involved in cell–cell adhesion. **(D)** GO enrichment analysis of molecular function. Highly enriched SC secreted proteins are involved in protein binding. Only the top seven cellular components, biological processes, and molecular functions have been shown here for vertical sizing. Additional details are listed in **Supplementary Tables S5–S7**. SC-CM, Schwann cell–conditioned media; GO, Gene Ontology; DAVID, Database for Annotation, Visualization, and Integrated Discovery.

30 min and 90°C for 10 s. IHC was then performed using an ImmPRESS^{TM} horseradish peroxidase (HRP) immunoglobulin G (peroxidase) Polymer Detection Kit (Vector Laboratories), as per the manufacturer's recommendations. After inactivation of endogenous peroxidases with 0.3% H_2O_2 and blocking with 2.5% horse serum, primary antibody followed by secondary antibodies was applied to the sections and revealed with DAB peroxidase (HRP) Substrate Kit (catalog number SK-4100, Vector Laboratories). Primary antibodies used are listed in **Supplementary Table S1**. Finally, TMA slides were counterstained with hematoxylin (Gill's formulation, Vector Laboratories), dehydrated, and cleared in xylene before mounting in Ultramount #4 mounting media (Thermo Fisher Scientific, Victoria, Australia). Following IHC staining, slides were scanned with a Leica Aperio AT2 Scanner (Leica Biosystems, Vista, CA, United States) (23).

Bioinformatics Analysis

UniProt was used to identify protein cellular localization. Functional clustering of the identified secreted proteins was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID, v6.8[1]) searching against the entire Homo sapiens genome (access date, March 25, 2020). DAVID classified the characteristic protein sets according to Gene Ontology (GO) terms for cellular compartments, biological processes, and molecular functions. DAVID was also used to recognize functional Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway categories. All the DAVID categories were ranked according to the number of proteins in each group and not with *p*-value. To investigate the potential associations between the identified proteins and diseases, the dataset was subjected to DAVID Genetic Association Database (GAD) analysis.

The cBio Cancer Genomics Portal[2] was used to determine the association of selected proteins with prognosis of PC using The Cancer Genome Atlas (TCGA) on pancreatic adenocarcinoma (access date, February 25, 2020). Our search was set to a total of 184 pancreatic adenocarcinoma samples (TCGA, PanCancer

[1]https://david.ncifcrf.gov/
[2]http://cbioportal.org

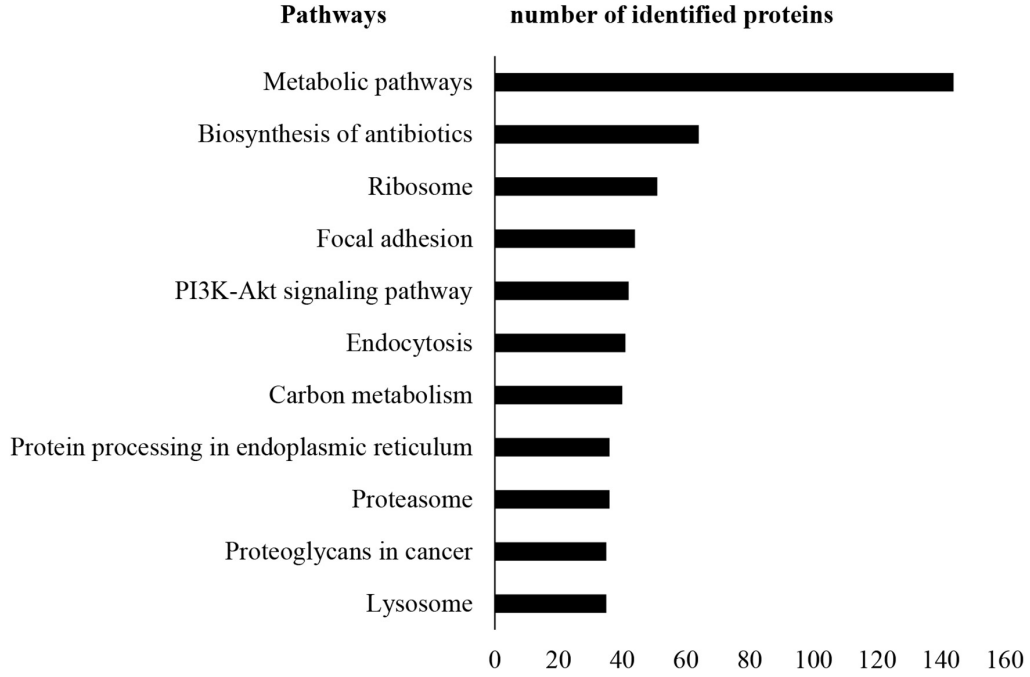


FIGURE 3 | KEGG pathway analysis. KEGG pathway enrichment analysis was performed using DAVID. Metabolic pathways were found as most enriched. Only the top 10 enriched pathways are shown here for vertical sizing. Additional details are listed in **Supplementary Table S8**. KEGG, Kyoto Encyclopedia of Genes and Genomes; DAVID, Database for Annotation, Visualization, and Integrated Discovery.

Atlas) and mRNA expression with a *z*-score threshold ±2.0. The role of selected candidates in cancer development was explored via the Cancer Hallmarks Analytics Tool, which allows organization and classification of cancer-related literature based on a text-mining analysis of 26 million PubMed abstracts (24).

Statistical Analysis

Statistical analysis was conducted using the GraphPad Prism software version 8.0 (GraphPad Software Inc., La Jolla, CA, United States). Statistical significance was determined by one-way analysis of variance (ANOVA). $p < 0.05$ was set the level of statistical significance. Data are presented as mean, with error bars representing the standard deviation. *p*-value is displayed as $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

RESULTS

Protein Map of SC Secretome and Pathway Analysis

A total of 13,796 unique peptides corresponding to 1,470 individual proteins were identified in two replicates, with a confidence corresponding to a false discovery rate <1% (**Figure 2A** and **Supplementary Table S2–S4**). Of the 1,470 proteins identified, 74% (1,084 proteins) were common across the two replicates. According to DAVID GO enrichment analysis of the cellular components, proteins localized in extracellular exosomes were highly enriched (692 proteins). Others were in cytoplasm (562 proteins), cytosol (542 proteins), nucleus (393 proteins), membrane (299 proteins), and nucleoplasm (257 proteins), as well as the extracellular space (198 proteins) (**Figure 2B** and **Supplementary Table S5**). To gain further insights into the enriched biological processes, common secreted proteins were subjected to GO enrichment analysis. The results showed that cell–cell adhesion was the most enriched (112 proteins) biological function. Other significant biological processes include translational initiation (72 proteins), oxidation–reduction process (65 proteins), and translation (61 proteins) (**Figure 2C** and **Supplementary Table S6**). GO enrichment analysis was also used to depict the molecular function of the identified proteins. Protein binding was the most enriched (794 proteins) molecular function. Other enriched molecular functions were mainly related to poly (A) RNA binding (260 proteins), cadherin binding (123 proteins), and ATP binding (120 proteins) (**Figure 2D** and **Supplementary Table S7**). KEGG pathway analysis was performed to map the important and representative pathways in human SCs using the DAVID resource, and the top 10 pathways based on enrichment were defined (**Figure 3** and **Supplementary Table S8**). Metabolic pathways were among the most enriched (144 proteins). Proteins from our list were also found to be potentially involved in focal adhesion (44 proteins), PI3K-Akt signaling pathway (42 proteins), endocytosis (41 proteins), protein processing in endoplasmic reticulum (36 proteins), and proteoglycans in cancer (35 proteins). Exploration of potential associations of the identified proteins with diseases using GAD resources revealed that approximately 21% (226 proteins) of the SC-secreted proteins were associated with cancer followed

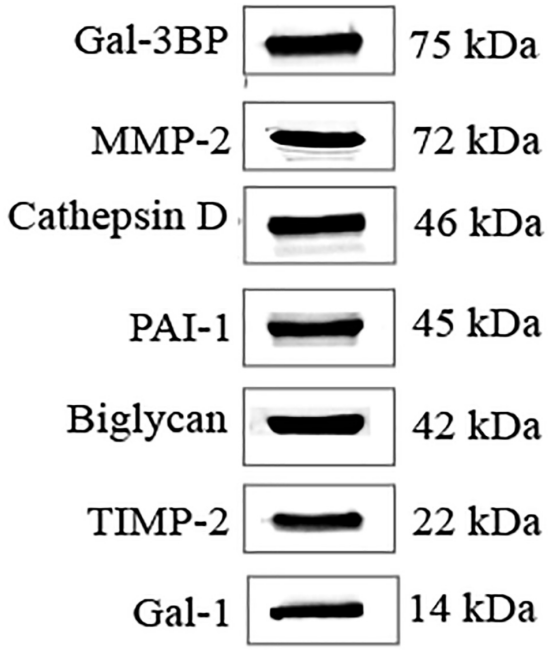


FIGURE 4 | WB validation of candidate molecules. WB analysis confirmed the presence of secreted candidate molecules in SC-CM. WB, western blot; SC-CM, Schwann cell–conditioned media; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1.

by neurological disorders (204 proteins) and infectious diseases (177 proteins). The remaining portion of our identified proteins was also associated with renal disease (113 proteins), aging (93 proteins), reproduction (73 proteins), and vision (58 proteins) (**Supplementary Figure S2** and **Supplementary Table S9**).

Protein Validation by WB

From the list of secretory proteins identified by LC-MS/MS, galectin-3–binding protein (Gal-3BP), matrix metalloproteinase-2 (MMP-2), cathepsin D, plasminogen activator inhibitor-1 (PAI-1), biglycan, tissue inhibitor of metalloproteinases-2 (TIMP-2), and galectin-1 (Gal-1) were validated by WB (**Figure 4**). We initially picked those proteins because they were known to be involved in tumor progression. These proteins are reported in **Table 2** with their identified peptide count, coverage, and observed function in PC progression. WB analyses of SC lysates for the identified proteins have also been provided in **Supplementary Figure S3**. Validated proteins with their Cancer Hallmarks Analytics Tool analysis results are presented in **Supplementary Figure S4**.

Impact of Targeting SC Secreted Proteins on PC Cell Growth

Significant increase of PC cells (PANC-1 and MIA PaCa-2) proliferation was observed in the presence of SC-CM compared with negative control (SF media, noted as SF) ($p < 0.05$), and the increment was similar to that in the positive control (serum-supplemented media, noted as FBS) (**Figure 5**), demonstrating that SC-CM stimulates PC cell proliferation. To determine if the proliferative effect of SC-CM was due to specific SC-secreted proteins, proliferation assay was performed in the presence of blocking antibodies against the WB-validated proteins. Blocking antibodies against Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, or Gal-1 caused a significantly decreased ($p < 0.05$) proliferation in PC cells compared to control (SC-CM only, without blocking antibody) (**Figures 5A,B**). No significant inhibitory effect on PC cell proliferation was observed when control media (SF and serum-supplemented media) was treated with blocking antibodies ($p > 0.05$) (**Supplementary Figure S5**).

Impact of Targeting SC Secreted Proteins on PC Cell Invasiveness

A significant increase in PC cell (both PANC-1 and MIA PaCa-2) invasion was observed in the presence of SC-CM compared with negative control (SF media, noted as SF) ($p < 0.05$), and the increment was similar to the positive control (serum-supplemented media, noted as FBS) (**Figure 6**). To specifically determine the SC-secreted proteins that induce increased invasion of PC cells, Transwell invasion assays were performed with blocking antibodies against the identified proteins. Among the blocking antibodies tested, neutralization of MMP-2, cathepsin D, PAI-1, and Gal-1 significantly ($p < 0.05$) decreased the invasion of PC cell lines compared to control

TABLE 2 | SCs secreted proteins whose targeting with antibodies resulted in a decreased PC cell proliferation and invasion.

No.	Accession no.	Protein name	Gene name	Unique peptides	MW (kDa)	Coverage (%)	Role in PC progression
1	Q08380	Gal-3BP	*LGALS3BP*	23	65.3	42.90	Proliferation
2	P08253	MMP-2	*MMP2*	40	73.8	61.81	Proliferation and invasion
3	P07339	Cathepsin D	*CTSD*	21	44.5	56.55	Proliferation and invasion
4	P05121	PAI-1	*SERPINE1*	41	45	62.68	Proliferation and invasion
5	P21810	Biglycan	*BGN*	18	41.6	53.53	Proliferation
6	P16035	TIMP-2	*TIMP2*	14	24.4	49.54	Proliferation
7	P09382	Gal-1	*LGALS1*	14	14.7	74.81	Proliferation and invasion

Commonly secreted proteins (between two replicates) identified using LC-MS/MS and validated by WB are presented here. Unique peptides: number of different peptides found by MS analyses; molecular weight: found by MS analyses; coverage: percentage of related protein found by MS analyses. LC-MS/MS, liquid chromatography–tandem mass spectrometry; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1.

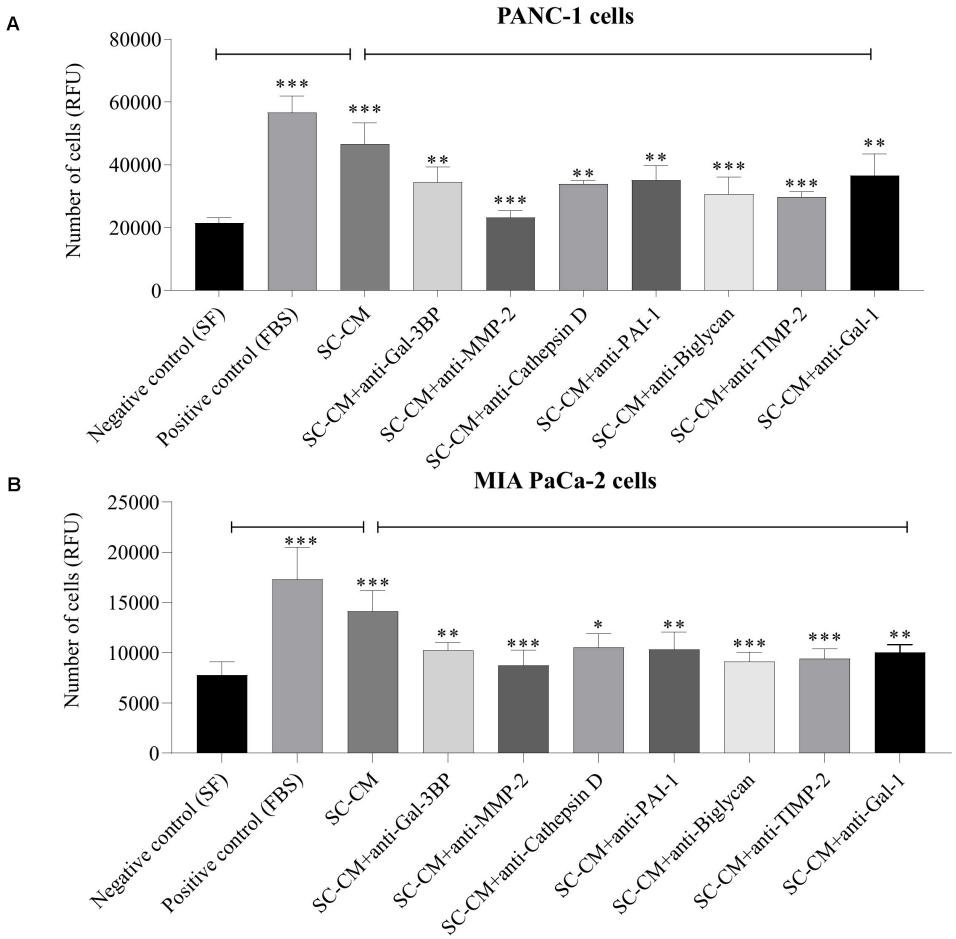


FIGURE 5 | *In vitro* blockage of SC-secreted proteins results in decrease of PC cell proliferation. The effects of SC-CM on PC cells (PANC-1 and MIA PaCa-2) proliferation were studied *in vitro* using resazurin-based proliferation assay. For both treatment and control, ∼5,000 cells were initially plated onto 96-well plate dishes and allowed to proliferate for 72 h in the presence of SC-CM or control media. Serum-free media was used as negative control, and serum-supplemented media was used as positive control. Cell number was quantified as described in section "Materials and Methods." A significantly greater number of PC cells were observed in presence of SC-CM compared to negative control ($p < 0.05$). To find out if the effect of increased proliferation of PC cells in presence of SC-CM was due to secreted proteins from SCs, proliferation assays were performed using blocking antibodies against the proteins of interest. Blocking antibodies were used at 6 μg/mL concentration. Significant decrease in cancer cell (both PANC-1 and MIA PaCa-2) proliferation was observed in presence of blocking antibodies against Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, and Gal-1 compared to SC-CM alone ($p < 0.05$). Representative data are shown from three independent experiments, performed in at least six replicates each. Statistical significance was confirmed by one-way ANOVA. The error bars represent the SD of the mean. *p*-values are displayed as $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. SC-CM, Schwann cell–conditioned media; PC, pancreatic cancer; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1; RFU, relative fluorescence unit.

(SC-CM only, without blocking antibody) (**Figures 6A,B**). In contrast, blocking Gal-3BP, biglycan, and TIMP-2 in SC-CM had no significant ($p > 0.05$) inhibitory effect on PC invasion. No significant decrease in PC cell invasiveness was observed when control media (both SF and serum-supplemented media) was treated with blocking antibodies ($p > 0.05$) (**Supplementary Figure S6**).

Immunohistochemical Detection of Identified Proteins in PC

The seven SC-secreted proteins having an effect on PC cell growth and invasion were investigated by IHC. The results revealed moderate to high levels of expression of these proteins in PC (**Supplementary Figures S7A–G**). Briefly, strong cytoplasmic immunoreactivity was observed in case of Gal-3 BP (**Supplementary Figure S7A**). In case of MMP-2, most malignant cells and surrounding stroma showed weak to moderate immunoreactivity (**Supplementary Figure S7B**). Cathepsin D was more strongly found in the PC cells and in SCs (**Supplementary Figure S7C**). In case of PAI-1, all cancer tissues were negative, and few tumor stroma were weakly positive (**Supplementary Figure S7D**). In case of biglycan, strong tumor and stromal immunoreactivity was observed (**Supplementary Figure S7E**). In case of TIMP-2, most cancer tissues were negative, whereas few malignant

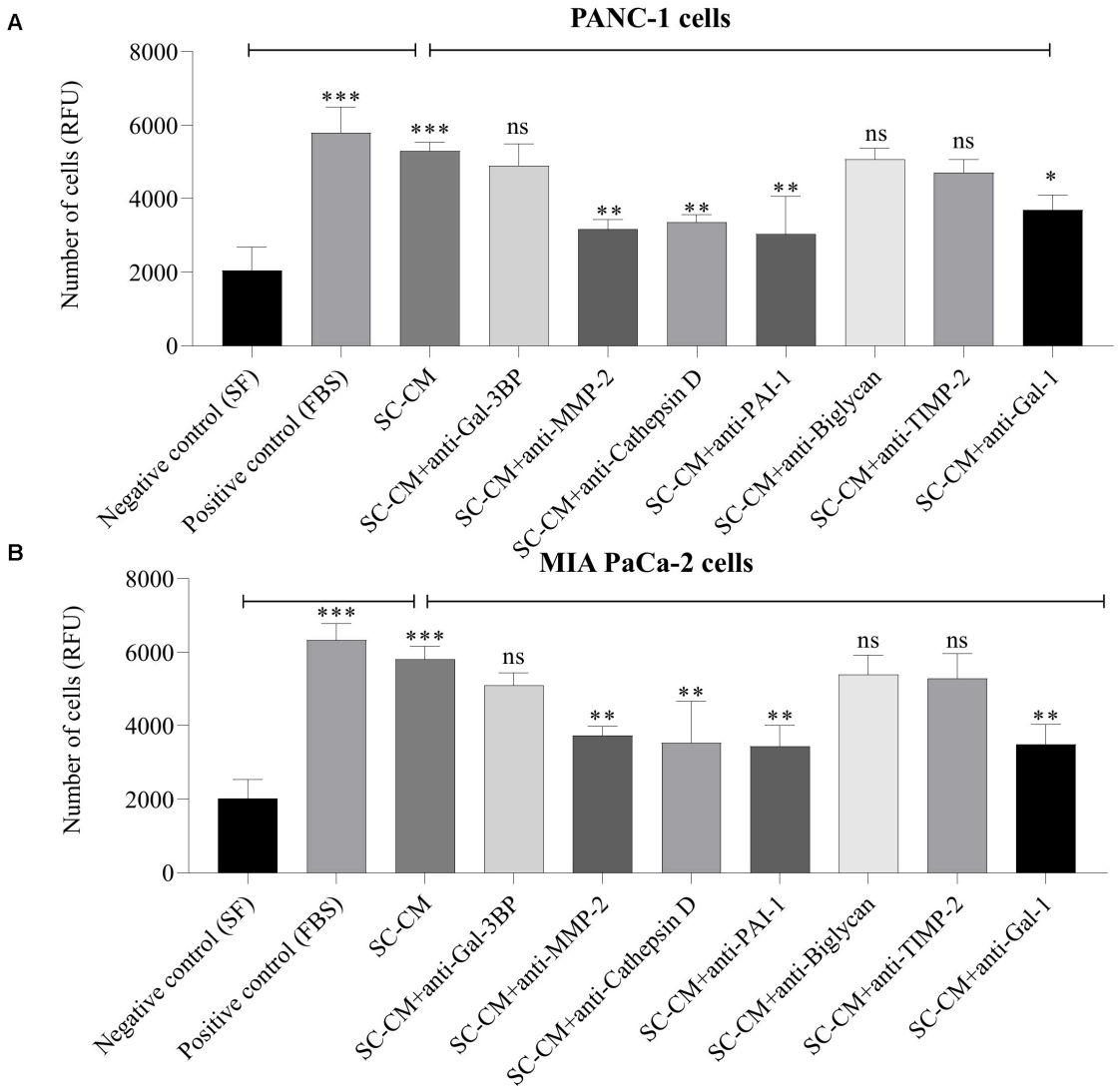


FIGURE 6 | *In vitro* blockage of SC-secreted proteins results in decreased invasion of PC cells. The effects of SC-CM on PC cells (PANC-1 and MIA PaCa-2) invasiveness were studied *in vitro* using Transwell invasion assay. For both treatment and control, cells (~60,000) were plated to the insert of a Transwell plate coated with collagen and allowed to invade for 24 h through a collagen-coated matrix toward SC-CM or control media as applicable. Serum-free media was used as negative control and serum-supplemented media was used as positive control. The non-invaded cells were removed from the top of the insert, and invaded cells were quantified using the QCM ECMatrix Cell Invasion Assay Kit described in section "Materials and Methods." A significantly greater number of invaded PC cells in presence of SC-CM was observed compared to negative control ($p < 0.05$). To determine if the proteins of interest were involved in the stimulation of PC cells by SCs, the invasion assay was performed with blocking antibodies. Among the tested blocking antibodies, neutralization of MMP-2, cathepsin D, PAI-1, and Gal-1 decreased the invasiveness of both types of PC cell line compared to SC-CM alone ($p < 0.05$). Blocking antibodies were used at 6 μg/mL concentration. Blocking Gal-3BP, biglycan, and TIMP-2 in SC-CM had no significant inhibitory effect on PC invasiveness ($p > 0.05$). Representative data are shown from three independent experiments, performed in at least three replicates each. Statistical significance was obtained by one-way ANOVA. The error bars represent the SD of the mean. *p*-values are displayed as $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. SC-CM, Schwann cell–conditioned media; PC, pancreatic cancer; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1; RFU, relative fluorescence unit; ns, non-significant.

cells displayed moderate cytoplasmic immunoreactivity (**Supplementary Figure S7F**). Gal-1 expression levels were mostly restricted to stroma of PC cells and in SCs (**Supplementary Figure S7G**). Positive labeling in IHC validates the presence of the identified proteins in the tumor microenvironment of PC. However, SCs are difficult to localize in the tumor microenvironment, and IHC cannot provide a clear demonstration that those proteins are released by SCs.

Prognostic Value of Protein Candidates

To explore the prognostic value of candidate proteins for which targeting with antibodies inhibited proliferation and invasion of PC cells, a meta-analysis at the mRNA was carried out

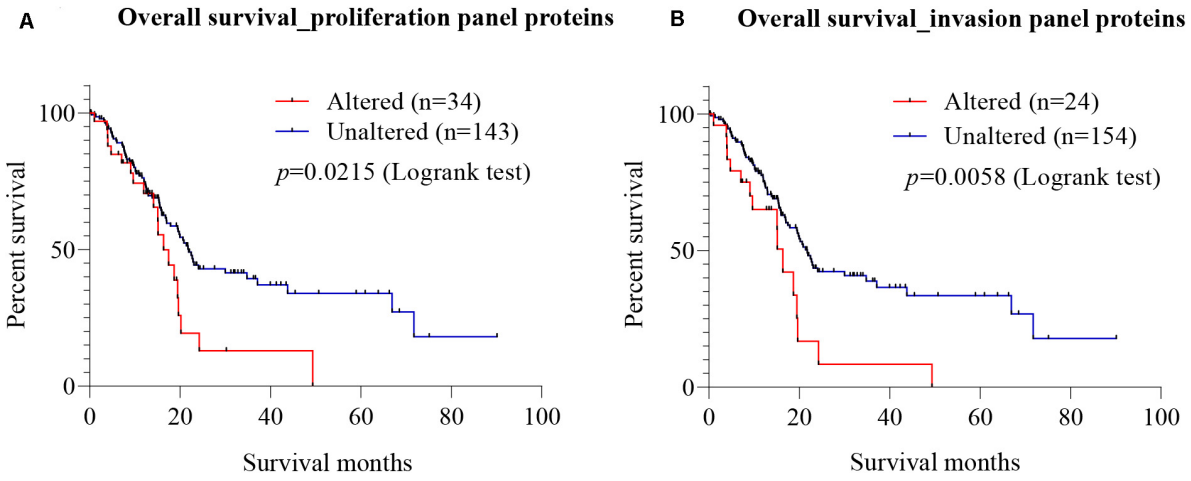


FIGURE 7 | Prognostic value of mRNA expression corresponding to the identified candidate proteins. Using cBioportal, Kaplan–Meier estimates of survival for PC patients with alterations in different gene combination were performed using The Cancer Genome Atlas (TCGA) on pancreatic adenocarcinoma. **(A)** Survival analysis showed that patients with alterations in the seven-gene proliferation panel (Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, and Gal-1) had worse overall survival than those without alterations (log-rank test $p = 0.0215$). **(B)** In the cases of invasion panel proteins (MMP-2, cathepsin D, PAI-1, and Gal-1), worse prognosis was also observed for patients with alterations in the four genes than those without alterations (log-rank test $p = 0.0058$). PC, pancreatic cancer; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1.

using TCGA on pancreatic adenocarcinoma, using the cBio Cancer Genomics Portal. Using Kaplan–Meier analysis of overall survival, based on the median mRNA expression levels, the prognostic values of the combined seven proteins involved in proliferation (Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, and Gal-1) (called proliferation protein panel) and four proteins involved in invasion (MMP-2, cathepsin D, PAI-1, and Gal-1) (called invasion protein panel) were investigated. The proliferation protein panel showed a significant association with poor prognosis (log-rank test *p*-value 0.0215), where 34 patients from a total of 177 showed alterations in this gene signature (**Figure 7A**). The invasion protein panel showed significant association with poor prognosis (log-rank test $p = 0.0058$), where 24 patients from a total of 178 showed alterations in this gene signature (**Figure 7B**). Overall, both protein panels may constitute a molecular signature for poor prognosis in PC. The data are useful in showing a prognostic relevance of the combination, but there is no demonstration that SCs are the major source for these proteins that may equally be produced by PC cells.

DISCUSSION

The present study has used proteomic analysis to define the secretome of SCs and has identified several proteins that can be targeted *in vitro* to inhibit growth and invasion of PC cells. These findings are summarized in **Figure 8**. In addition, several of the identified proteins were shown to contribute to PC cell growth and invasion and may constitute future therapeutic targets.

The LC-MS/MS–based proteomic analysis that has been implemented in this study has enabled the identification of 1,084 SC-secreted proteins. Bioinformatics analysis was performed. In GO analysis, the "extracellular exosome" was found to be the most represented localization, and that is in accordance with the fact that the exosome compartment is a key part of the secretome. Cell–cell adhesion was the most enriched biological function, and this is in line with the supportive role of SC in nerves (25) and their role in promoting the nerve-cancer cell interaction (11). Strikingly, GO analysis also revealed that 73% of the identified proteins possessed molecular functions related to protein binding. Gal-3 BP or insulin-like growth factor BP (IGFBP), for instance, is well described for its binding activities that contribute to the regulation of cell growth (26, 27). The SC secretome was also enriched in proteins involved in molecular functions related to "catalytic activity," which may play important roles in cancer progression. For instance, cathepsin D is a catalytic protein that stimulates cancer cell proliferation and tumor angiogenesis and can also provide protection against tumor apoptosis (28). Similarly, other significantly enriched molecular functions including receptor binding, fibroblast growth factor binding, and platelet-derived growth factor binding are all pertinent to tumorigenesis and metastasis. Collectively, these findings indicate that the protein signatures identified in the SC secretome match molecular networks and biological processes associated with tumor progression. In future studies, it would be of interest to investigate the proteome of SCs when stimulated by PC cells to further delineate the crosstalk between PC cells and SC in the tumor microenvironment.

Some of the identified proteins were validated by WB, and blocking antibodies were used to test the effect of their inhibition on PC cell proliferation and invasion *in vitro*. We observed that blocking antibodies against Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, and Gal-1 inhibited the proliferation of PC cells induced by SC-CM. Additionally, blocking MMP-2, cathepsin D, PAI-1, and Gal-1also reduced cancer cell invasion.

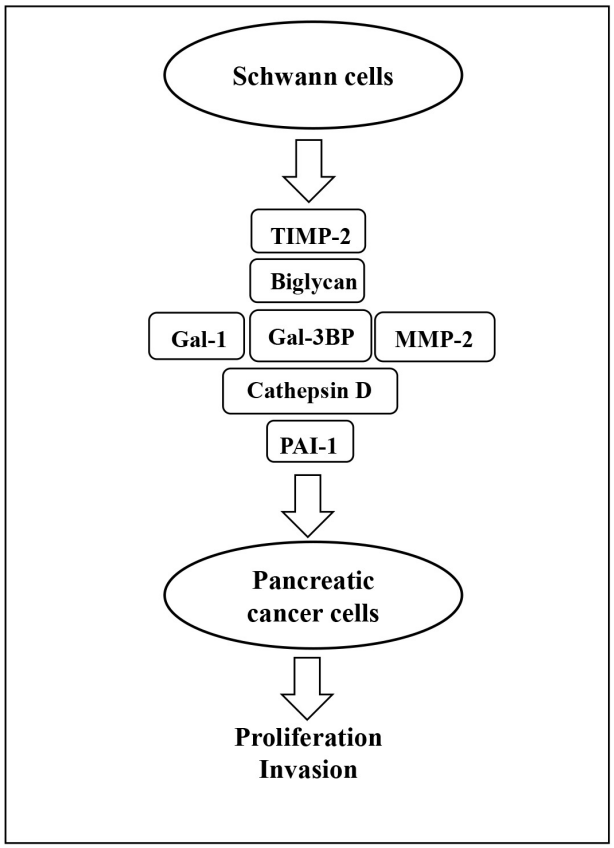


FIGURE 8 | Schematic representation of SC-stimulation of PC cells. Gal-3BP, MMP-2, cathepsin D, PAI-1, biglycan, TIMP-2, and Gal-1 are the SC-secreted proteins that have been identified as potential promoters of PC cell proliferation and invasion in this study. SC, Schwann cell; PC, pancreatic cancer; Gal-3BP, galectin-3 binding protein; MMP-2, matrix metalloproteinase-2; PAI-1, plasminogen activator inhibitor-1; TIMP-2, tissue inhibitor of metalloproteinases-2; Gal-1, galectin-1.

The potential role of each of these proteins in PC cancer is discussed below.

Galectin-3 BP, also known as tumor-associated antigen 90K, is a large oligomeric heavily glycosylated and secreted protein (26). It is a binding partner of Gal-1 and Gal-3, which promote integrin-mediated cell adhesion, and significantly elevated expression of Gal-3 BP in the serum or tumor tissues is associated with poor prognosis in a variety of malignancies including breast cancer (29), lung cancer (30), and PC (31). Additionally, Gal-3 BP is involved in the promotion of integrin-mediated tumor cell adhesion to the ECM proteins in colon cancer (32), breast cancer (33), and the formation of metastasis in lung cancer (34). Our study reveals that targeting Gal-3 BP may prevent the stimulatory effect of SCs in PC cancer cells. In addition, Gal-1, a binding target of Gal-3 BP, has also been identified in our study.

Galectin-1 is a dimeric carbohydrate BP that facilitates the malignant cellular activities by cross-linking glycoproteins (35). It has been reported to play a role in cell invasion of several tumor types, including pancreatic (36), lung (37), and epithelial ovarian tumors (38). Knockdown of this protein can decrease the invasiveness of cancer cells in cervical cancer (39) and oral squamous cell carcinomas (40). Gal-1 has been reported to promote cancer cell invasion by enhancing the expression and enzymatic activities of MMP-2 and MMP-9 (40). It also appears to promote epithelial–mesenchymal transition in lung cancer cell lines (41), and our study points to the role of this protein in the SC-induced stimulation of PC cell growth and invasion.

Matrix metalloproteinase-2, a zinc-dependent endopeptidase, has been implicated in the malignant potential of tumor cells, because of its ability to degrade ECM proteins (42). MMP-2 is associated with the development of desmoplastic reaction in PC (43), and downregulation of MMP-2 reduces PC cell migration (44) and invasion (45). Our study reveals that MMP-2 is a potential mediator of the stimulatory role of SCs in PC cell proliferation and invasion. Interestingly, TIMP-2, tissue inhibitor of metalloproteinases (TIMPs) family, has also been identified in our study. TIMP-2 was described to decrease cell proliferation and migration *in vitro* via the inhibition of MMPs (46). However, it is associated with poor patient outcomes in cancers including gastric (47), renal (48), and oral squamous cell cancers (49). In lung cancer, TIMP-2 has been reported to inhibit tumor growth by promoting an antitumoral transcriptional profile both *in vitro* and *in vivo* (50). High expression of TIMP-2 has been shown to correlate with adverse prognosis in breast cancer (51). Our study indicates that MMP-2 and TIMP-2 are both released by SC and that they can contribute to the stimulation of PC cell growth and invasion. Similarly, PAI-1, another protease inhibitor also known as serpin (52), has also been identified in our study. It has been reported to inhibit proliferation of hepatocellular (53) and prostate cancer cell growth (54). In ovarian cancer, PAI-1 facilitates cell growth and inhibits apoptosis (55). Overexpression of PAI-1 inhibited cell migration and invasion in PC (56). Our study shows that PAI-1 is secreted by SC and can stimulate PC cell proliferation and invasion.

Cathepsin D is a secreted aspartic protease, which when highly expressed is associated with unfavorable clinical outcomes in patients with PC (57). It has been shown that cathepsin D expression can accelerate the metastatic spread of PC by upregulation of S100P (58). Combination of cathepsin D with CA-19-9 and MMP-7 has been reported to be an important panel of markers for screening PC (59). High cathepsin D expression in PC has been shown to decrease the effectiveness of adjuvant gemcitabine (60). The present study reveals the potential role of cathepsin D as a promoter of PC cell proliferation and invasion induced by SC.

Biglycan is a leucine-rich proteoglycan whose overexpression is related to enhanced angiogenesis and tumor invasion (61). Correlations of biglycan expression with aggressive clinicopathological features and poor survival in human cancers such as pancreatic adenocarcinoma (62), colorectal cancer (61), and gastric cancer (63) have been reported. High biglycan expression has been shown to promote invasiveness of melanoma cells (64). It has been reported to promote tumor invasion, migration, and metastasis of gastric cancer cells both *in vitro* and *in vivo* through activating the FAK signaling pathway (63). Our

results suggest that biglycan is secreted by SC and can stimulate PC cell proliferation.

Our proteomic analysis also identified several other proteins of interest that, although they were not tested in functional analysis, may still be important in the stimulation PC cells by SCs. This is, for instance, the case of IGFBPs (IGFBP-2, IGFBP-4, IGFBP-5, IGFBP-6, IGFBP-7) and transforming growth factors (TGF-β1, TGF-β2, TGF-BI). Further studies are warranted to determine the possible involvement of these proteins in cancer PC progression and their potential value as therapeutic targets.

CONCLUSION

In summary, this proteomic and functional analysis identified a number of SC-secreted proteins that seem to be involved in the proliferation and invasiveness of PC cancer cells induced by SC. Further preclinical studies *in vivo* are warranted to determine whether these proteins may become new targets for therapeutic intervention in PC.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: the Mass Spectrometry Interactive Virtual Environment (MassIVE) database with the dataset identifier MSV000084303 (https://massive.ucsd.edu/ProteoSAFe).

AUTHOR CONTRIBUTIONS

HH and AF designed the study and wrote the manuscript. AF and XL carried out the experiments. AF analyzed the data. XL, NG, SF, MJ, FG, CJ, DH, PT, and PJ edited the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was supported by the Maitland Cancer Appeal Committee (NSW Australia), the Hunter Medical Research Institute (HMRI), and the Hunter Cancer Research Alliance (HCRA).

ACKNOWLEDGMENTS

The authors thank the University of Newcastle, and the funding bodies for supporting this research and Kristen McEwan as well as Mark Marshland for providing excellent research assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01601/full#supplementary-material

REFERENCES

1. Are C, Chowdhury S, Ahmad H, Ravipati A, Song T, Shrikandhe S, et al. Predictive global trends in the incidence and mortality of pancreatic cancer based on geographic location, socio-economic status, and demographic shift. *J Surg Oncol*. (2016) 114:736–42. doi: 10.1002/jso.24410
2. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res*. (2014) 74:2913–21. doi: 10.1158/0008-5472.Can-14-0155
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. (2018) 68:7–30. doi: 10.3322/caac.21442
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. (2017) 67:7–30. doi: 10.3322/caac.21387
5. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med*. (2014) 371:1039–49. doi: 10.1056/NEJMra1404198
6. Takahashi Y, Sipp D, Enomoto H. Tissue interactions in neural crest cell development and disease. *Science*. (2013) 341:860–3. doi: 10.1126/science.1230717
7. Jessen KR. Glial cells. *Int J Biochem Cell Biol*. (2004) 36:1861–7. doi: 10.1016/j.biocel.2004.02.023
8. Kaplan S, Odaci E, Unal B, Sahin B, Fornaro M. Chapter 2: development of the peripheral nerve. *Int Rev Neurobiol*. (2009) 87:9–26. doi: 10.1016/s0074-7742(09)87002-5
9. Lehmann HC, Hoke A. Use of engineered Schwann cells in peripheral neuropathy: hopes and hazards. *Brain Res*. (2016) 1638(Pt A):97–104. doi: 10.1016/j.brainres.2015.10.040
10. Demir IE, Boldis A, Pfitzinger PL, Teller S, Brunner E, Klose N, et al. Investigation of Schwann cells at neoplastic cell sites before the onset of cancer invasion. *J Natl Cancer Inst*. (2014) 106:dju184. doi: 10.1093/jnci/dju184
11. Deborde S, Omelchenko T, Lyubchik A, Zhou Y, He S, McNamara WF, et al. Schwann cells induce cancer cell dispersion and invasion. *J Clin Investig*. (2016) 126:1538–54. doi: 10.1172/jci82658
12. Fujii-Nishimura Y, Yamazaki K, Masugi Y, Douguchi J, Kurebayashi Y, Kubota N, et al. Mesenchymal-epithelial transition of pancreatic cancer cells at perineural invasion sites is induced by Schwann cells. *Pathol Int*. (2018) 68:214–23. doi: 10.1111/pin.12641
13. Sroka IC, Chopra H, Das L, Gard JMC, Nagle RB, Cress AE. Schwann cells increase prostate and pancreatic tumor cell invasion using laminin binding A6 integrin. *J Cell Biochem*. (2016) 117:491–9. doi: 10.1002/jcb.25300
14. Zhou Y, Shurin GV, Zhong H, Bunimovich YL, Han B, Shurin MR. Schwann cells augment cell spreading and metastasis of lung cancer. *Cancer Res*. (2018) 78:5927–39. doi: 10.1158/0008-5472.CAN-18-1702
15. Salvo E, Saraithong P, Curtin JG, Janal MN, Ye Y. Reciprocal interactions between cancer and Schwann cells contribute to oral cancer progression and pain. *Heliyon*. (2019) 5:e01223. doi: 10.1016/j.heliyon.2019.e01223
16. Huang T, Fan Q, Wang Y, Cui Y, Wang Z, Yang L, et al. Schwann cell-derived CCL2 promotes the perineural invasion of cervical cancer. *Front Oncol*. (2020) 10:19. doi: 10.3389/fonc.2020.00019
17. Demir IE, Tieftrunk E, Schorn S, Saricaoglu OC, Pfitzinger PL, Teller S, et al. Activated Schwann cells in pancreatic cancer are linked to analgesia via suppression of spinal astroglia and microglia. *Gut*. (2016) 65:1001–14. doi: 10.1136/gutjnl-2015-309784
18. Su D, Guo X, Huang L, Ye H, Li Z, Lin L, et al. Tumor-neuroglia interaction promotes pancreatic cancer metastasis. *Theranostics*. (2020) 10:5029–47. doi: 10.7150/thno.42440
19. Roger E, Martel S, Bertrand-Chapel A, Depollier A, Chuvin N, Pommier RM, et al. Schwann cells support oncogenic potential of pancreatic cancer cells through TGFβ signaling. *Cell Death Dis*. (2019) 10:886. doi: 10.1038/s41419-019-2116-x

20. Na'ara S, Amit M, Gil Z. L1CAM induces perineural invasion of pancreas cancer cells by upregulation of metalloproteinase expression. *Oncogene*. (2019) 38:596–608. doi: 10.1038/s41388-018-0458-y
21. Ferdoushi A, Li X, Jamaluddin MFB, Hondermarck H. Proteomic profile of human Schwann cells. *Proteomics*. (2020) 20:1900294. doi: 10.1002/pmic.201900294
22. Faulkner S, Jobling P, Rowe CW, Rodrigues Oliveira SM, Roselli S, Thorne RF, et al. Neurotrophin receptors TrkA, p75(NTR), and sortilin are increased and targetable in thyroid cancer. *Am J Pathol*. (2018) 188:229–41. doi: 10.1016/j.ajpath.2017.09.008
23. Gao F, Griffin N, Faulkner S, Rowe CW, Williams L, Roselli S, et al. The neurotrophic tyrosine kinase receptor TrkA and its ligand NGF are increased in squamous cell carcinomas of the lung. *Sci Rep*. (2018) 8:8135. doi: 10.1038/s41598-018-26408-2
24. Baker S, Ali I, Silins I, Pyysalo S, Guo Y, Hogberg J, et al. Cancer hallmarks analytics tool (CHAT): a text mining approach to organize and evaluate scientific literature on cancer. *Bioinformatics*. (2017) 33:3973–81. doi: 10.1093/bioinformatics/btx454
25. Corfas G, Velardez MO, Ko C-P, Ratner N, Peles E. Mechanisms and roles of axon-schwann cell interactions. *J Neurosci*. (2004) 24:9250. doi: 10.1523/JNEUROSCI.3649-04.2004
26. Koths K, Taylor E, Halenbeck R, Casipit C, Wang A. Cloning and characterization of a human Mac-2-binding protein, a new member of the superfamily defined by the macrophage scavenger receptor cysteine-rich domain. *J Biol Chem*. (1993) 268:14245–9.
27. Forbes BE, McCarthy P, Norton RS. Insulin-like growth factor binding proteins: a structural perspective. *Front Endocrinol*. (2012) 3:38. doi: 10.3389/fendo.2012.00038
28. Berchem G, Glondu M, Gleizes M, Brouillet JP, Vignon F, Garcia M, et al. Cathepsin-D affects multiple tumor progression steps in vivo: proliferation, angiogenesis and apoptosis. *Oncogene*. (2002) 21:5951–5. doi: 10.1038/sj.onc.1205745
29. Tinari N, Lattanzio R, Querzoli P, Natoli C, Grassadonia A, Alberti S, et al. High expression of 90K (Mac-2 BP) is associated with poor survival in node-negative breast cancer patients not receiving adjuvant systemic therapies. *Int J Cancer*. (2009) 124:333–8. doi: 10.1002/ijc.23970
30. Marchetti A, Tinari N, Buttitta F, Chella A, Angeletti CA, Sacco R, et al. Expression of 90K (Mac-2 BP) correlates with distant metastasis and predicts survival in stage I non-small cell lung cancer patients. *Cancer Res*. (2002) 62:2535–9.
31. Kunzli BM, Berberat PO, Zhu ZW, Martignoni M, Kleeff J, Tempia-Caliera AA, et al. Influences of the lysosomal associated membrane proteins (Lamp-1, Lamp-2) and Mac-2 binding protein (Mac-2-BP) on the prognosis of pancreatic carcinoma. *Cancer*. (2002) 94:228–39. doi: 10.1002/cncr.10162
32. Ulmer TA, Keeler V, Loh L, Chibbar R, Torlakovic E, Andre S, et al. Tumor-associated antigen 90K/Mac-2-binding protein: possible role in colon cancer. *J Cell Biochem*. (2006) 98:1351–66. doi: 10.1002/jcb.20784
33. Noma N, Simizu S, Kambayashi Y, Kabe Y, Suematsu M, Umezawa K. Involvement of NF-kappaB-mediated expression of galectin-3-binding protein in TNF-alpha-induced breast cancer cell adhesion. *Oncol Rep*. (2012) 27:2080–4. doi: 10.3892/or.2012.1733
34. Ozaki Y, Kontani K, Teramoto K, Fujita T, Tezuka N, Sawai S, et al. Involvement of 90K/Mac-2 binding protein in cancer metastases by increased cellular adhesiveness in lung cancer. *Oncol Rep*. (2004) 12:1071–7.
35. Cousin JM, Cloninger MJ. The role of galectin-1 in cancer progression, and synthetic multivalent systems for the study of galectin-1. *Int J Mol Sci*. (2016) 17:1566. doi: 10.3390/ijms17091566
36. Xue X, Lu Z, Tang D, Yao J, An Y, Wu J, et al. Galectin-1 secreted by activated stellate cells in pancreatic ductal adenocarcinoma stroma promotes proliferation and invasion of pancreatic cancer cells: an in vitro study on the microenvironment of pancreatic ductal adenocarcinoma. *Pancreas*. (2011) 40:832–9. doi: 10.1097/MPA.0b013e318217945e
37. Chung L-Y, Tang S-J, Sun G-H, Chou T-Y, Yeh T-S, Yu S-L, et al. Galectin-1 promotes lung cancer progression and chemoresistance by upregulating p38 MAPK, ERK, and cyclooxygenase-2. *Clin Cancer Res*. (2012) 18:4037–47. doi: 10.1158/1078-0432.ccr-11-3348
38. Kim HJ, Jeon HK, Cho YJ, Park YA, Choi JJ, Do IG, et al. High galectin-1 expression correlates with poor prognosis and is involved in epithelial ovarian cancer proliferation and invasion. *Eur J Cancer*. (2012) 48:1914–21. doi: 10.1016/j.ejca.2012.02.005
39. Kim HJ, Do IG, Jeon HK, Cho YJ, Park YA, Choi JJ, et al. Galectin 1 expression is associated with tumor invasion and metastasis in stage IB to IIA cervical cancer. *Hum Pathol*. (2013) 44:62–8. doi: 10.1016/j.humpath.2012.04.010
40. Wu M-H, Hong T-M, Cheng H-W, Pan S-H, Liang Y-R, Hong H-C, et al. Galectin-1-mediated tumor invasion and metastasis, up-regulated matrix metalloproteinase expression, and reorganized actin cytoskeletons. *Mol Cancer Res*. (2009) 7:311. doi: 10.1158/1541-7786.MCR-08-0297
41. Hsu Y-L, Wu C-Y, Hung J-Y, Lin Y-S, Huang M-S, Kuo P-L. Galectin-1 promotes lung cancer tumor metastasis by potentiating integrin 6 4 and Notch1/Jagged2 signaling pathway. *Carcinogenesis*. (2013) 34:1370–81. doi: 10.1093/carcin/bgt040
42. Nelson AR, Fingleton B, Rothenberg ML, Matrisian LM. Matrix metalloproteinases: biologic activity and clinical implications. *J Clin Oncol*. (2000) 18:1135–49. doi: 10.1200/jco.2000.18.5.1135
43. Ellenrieder V, Alber B, Lacher U, Hendler SF, Menke A, Boeck W, et al. Role of MT-MMPs and MMP-2 in pancreatic cancer progression. *Int J Cancer*. (2000) 85:14–20. doi: 10.1002/(sici)1097-0215(20000101)85:13.0.co;2-o
44. Zou Y, Li J, Chen Z, Li X, Zheng S, Yi D, et al. miR-29c suppresses pancreatic cancer liver metastasis in an orthotopic implantation model in nude mice and affects survival in pancreatic cancer patients. *Carcinogenesis*. (2015) 36:676–84. doi: 10.1093/carcin/bgv027
45. Layton T, Stalens C, Gunderson F, Goodison S, Silletti S. Syk tyrosine kinase acts as a pancreatic adenocarcinoma tumor suppressor by regulating cellular growth and invasion. *Am J Pathol*. (2009) 175:2625–36. doi: 10.2353/ajpath.2009.090543
46. Wang W, Li D, Xiang L, Lv M, Tao L, Ni T, et al. TIMP-2 inhibits metastasis and predicts prognosis of colorectal cancer via regulating MMP-9. *Cell Adh Migr*. (2019) 13:273–84. doi: 10.1080/19336918.2019.1639303
47. Hsu KW, Fang WL, Huang KH, Huang TT, Lee HC, Hsieh RH, et al. Notch1 pathway-mediated microRNA-151-5p promotes gastric cancer progression. *Oncotarget*. (2016) 7:38036–51. doi: 10.18632/oncotarget.9342
48. Kawata N, Nagane Y, Hirakata H, Ichinose T, Okada Y, Yamaguchi K, et al. Significant relationship of matrix metalloproteinase 9 with nuclear grade and prognostic impact of tissue inhibitor of metalloproteinase 2 for incidental clear cell renal cell carcinoma. *Urology*. (2007) 69:1049–53. doi: 10.1016/j.urology.2007.02.044
49. Katayama A, Bandoh N, Kishibe K, Takahara M, Ogino T, Nonaka S, et al. Expressions of matrix metalloproteinases in early-stage oral squamous cell carcinoma as predictive indicators for tumor metastases and prognosis. *Clin Cancer Res*. (2004) 10:634–40. doi: 10.1158/1078-0432.ccr-0864-02
50. Bourboulia D, Han H, Jensen-Taubman S, Gavil N, Isaac B, Wei B, et al. TIMP-2 modulates cancer cell transcriptional profile and enhances E-cadherin/beta-catenin complex expression in A549 lung cancer cells. *Oncotarget*. (2013) 4:166–76. doi: 10.18632/oncotarget.801
51. Remacle A, McCarthy K, Noel A, Maguire T, McDermott E, O'Higgins N, et al. High levels of TIMP-2 correlate with adverse prognosis in breast cancer. *Int J Cancer*. (2000) 89:118–21. doi: 10.1002/(sici)1097-0215(20000320)89:23.0.co;2-8
52. Fang H, Placencio VR, DeClerck YA. Protumorigenic activity of plasminogen activator inhibitor-1 through an antiapoptotic function. *J Natl Cancer Inst*. (2012) 104:1470–84. doi: 10.1093/jnci/djs377
53. Morita Y, Hayashi Y, Kanamaru T, Itoh T, Suzuki S, Yamamoto M, et al. Inhibitory role of plasminogen activator inhibitor-1 in invasion and proliferation of HLE hepatocellular carcinoma cells. *Jpn J Cancer Res*. (1999) 90:747–52. doi: 10.1111/j.1349-7006.1999.tb00810.x
54. Chen SC, Henry DO, Reczek PR, Wong MK. Plasminogen activator inhibitor-1 inhibits prostate tumor growth through endothelial apoptosis. *Mol Cancer Ther*. (2008) 7:1227–36. doi: 10.1158/1535-7163.Mct-08-0051
55. Mashiko S, Kitatani K, Toyoshima M, Ichimura A, Dan T, Usui T, et al. Inhibition of plasminogen activator inhibitor-1 is a potential therapeutic strategy in ovarian cancer. *Cancer Biol Ther*. (2015) 16:253–60. doi: 10.1080/15384047.2014.1001271
56. Inoue M, Sawada T, Uchima Y, Kimura K, Nishihara T, Tanaka H, et al. Plasminogen activator inhibitor-1 (PAI-1) gene transfection inhibits the liver metastasis of pancreatic cancer by preventing angiogenesis. *Oncol Rep*. (2005) 14:1445–51.

57. Niedergethmann M, Hildenbrand R, Wolf G, Verbeke CS, Richter A, Post S. Angiogenesis and cathepsin expression are prognostic factors in pancreatic adenocarcinoma after curative resection. *Int J Pancreatol*. (2000) 28:31–9. doi: 10.1385/ijgc:28:1:31
58. Whiteman HJ, Weeks ME, Dowen SE, Barry S, Timms JF, Lemoine NR, et al. The role of S100P in the invasion of pancreatic cancer cells is mediated through cytoskeletal changes and regulation of cathepsin D. *Cancer Res*. (2007) 67:8633–42. doi: 10.1158/0008-5472.Can-07-0545
59. Park HD, Kang ES, Kim JW, Lee KT, Lee KH, Park YS, et al. Serum CA19-9, cathepsin D, and matrix metalloproteinase-7 as a diagnostic panel for pancreatic ductal adenocarcinoma. *Proteomics*. (2012) 12:3590–7. doi: 10.1002/pmic.201200101
60. Mahajan UM, Goni E, Langhoff E, Li Q, Costello E, Greenhalf W, et al. Cathepsin D expression and gemcitabine resistance in pancreatic cancer. *JNCI Cancer Spectr*. (2019) 4:kz060. doi: 10.1093/jncics/pkz060
61. Gu X, Ma Y, Xiao J, Zheng H, Song C, Gong Y, et al. Up-regulated biglycan expression correlates with the malignancy in human colorectal cancers. *Clin Exp Med*. (2012) 12:195–9. doi: 10.1007/s10238-011-0155-4
62. Aprile G, Avellini C, Reni M, Mazzer M, Foltran L, Rossi D, et al. Biglycan expression and clinical outcome in patients with pancreatic adenocarcinoma. *Tumor Biol*. (2013) 34:131–7. doi: 10.1007/s13277-012-0520-2
63. Hu L, Duan Y-T, Li J-F, Su L-P, Yan M, Zhu Z-G, et al. Biglycan enhances gastric cancer invasion by activating FAK signaling pathway. *Oncotarget*. (2014) 5:1885–96. doi: 10.18632/oncotarget.1871
64. Andrlova H, Mastroianni J, Madl J, Kern JS, Melchinger W, Dierbach H, et al. Biglycan expression in the melanoma microenvironment promotes invasiveness via increased tissue stiffness inducing integrin-beta1 expression. *Oncotarget*. (2017) 8:42901–16. doi: 10.18632/oncotarget.17160

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 © 2020 Ferdoushi, Li, Griffin, Faulkner, Jamaluddin, Gao, Jiang, van Helden, Tanwar, Jobling and Hondermarck. 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 Oncology

ORIGINAL RESEARCH
published: 25 August 2020
doi: 10.3389/fonc.2020.01494

Mutant Proteomics of Lung Adenocarcinomas Harboring Different *EGFR* Mutations

Toshihide Nishimura[1,2*†], *Ákos Végvári*[3†], *Haruhiko Nakamura*[2], *Harubumi Kato*[4,5] *and Hisashi Saji*[2]

[1] Department of Translational Medicine Informatics, St. Marianna University School of Medicine, Kawasaki, Japan, [2] Department of Chest Surgery, St. Marianna University School of Medicine, Kawasaki, Japan, [3] Division of Chemistry I, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden, [4] Division of Thoracic and Thyroid Surgery, Tokyo Medical University, Tokyo, Japan, [5] Research Institute of Health and Welfare Sciences, Graduate School, International University of Health and Welfare, Tokyo, Japan

Epidermal growth factor receptor *EGFR* major driver mutations may affect downstream molecular networks and pathways, which would influence treatment outcomes of non-small cell lung cancer (NSCLC). This study aimed to unveil profiles of mutant proteins expressed in lung adenocarcinomas of 36 patients harboring representative driver *EGFR* mutations (Ex19del, nine; L858R, nine; no Ex19del/L858R, 18). Surprisingly, the orthogonal partial least squares discriminant analysis performed for identified mutant proteins demonstrated the profound differences in distance among the different *EGFR* mutation groups, suggesting that cancer cells harboring L858R or Ex19del emerge from cellular origins different from L858R/Ex19del-negative cells. Weighted gene coexpression network analysis, together with over-representative analysis, identified 18 coexpressed modules and their eigen proteins. Pathways enriched differentially for both the L858R and Ex19del mutations included carboxylic acid metabolic process, cell cycle, developmental biology, cellular responses to stress, mitotic prophase, cell proliferation, growth, epithelial to mesenchymal transition (EMT), and immune system. The IPA causal network analysis identified the highly activated networks of *PARPBP*, *HOXA1*, and *APH1* under the L858R mutation, whereas those of *ASGR1*, *APEX1*, *BUB1*, and *MAPK10* were highly activated under the Ex19del mutation. Interestingly, the downregulated causal network of osimertinib intervention showed the highest significance in overlap *p*-value among most causal networks predicted under the L858R mutation. We also identified the causal network of MAPK interacting serine/threonine kinase 1/2 (*MNK1/2*) highly activated differentially under the L858R mutation. Tumor-suppressor *AMOT*, a component of the Hippo pathways, was highly inhibited commonly under both L858R and Ex19del mutations. Our results could identify disease-related protein molecular networks from the landscape of single amino acid variants. Our findings may help identify potential therapeutic targets and develop therapeutic strategies to improve patient outcomes.

Keywords: mutant proteomics, lung adenocarcinoma, EGFR mutation, WGCNA, causal network analysis, disease-related molecular networks, mass spectrometry

OPEN ACCESS

Edited by:
Harsha Gowda, The University of Queensland, Australia

Reviewed by:
Francesco Grignani, University of Perugia, Italy
Rafael Rosell, Catalan Institute of Oncology, Spain

****Correspondence:***
Toshihide Nishimura
t-nisimura@marianna-u.ac.jp

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *29 April 2020*
Accepted: *13 July 2020*
Published: *25 August 2020*

Citation:
Nishimura T, Végvári Á, Nakamura H, Kato H and Saji H (2020) Mutant Proteomics of Lung Adenocarcinomas Harboring Different EGFR Mutations. Front. Oncol. 10:1494. doi: 10.3389/fonc.2020.01494

HIGHLIGHTS

- The first study to perform mutant proteomic analysis of clinical tissue specimens obtained from patients of lung adenocarcinoma with *EGFR* oncogenic driver mutations.
- Surprisingly, the OPLS discriminant analysis revealed profound differences among the profiles of mutant proteins identified under the different *EGFR* mutation statuses, which were never seen before.
- Weighted gene coexpression network analysis (WGCNA) screened by the over-representative test identified 18 significant network modules under the respective *EGFR* mutation statuses.
- Interestingly, the downregulated causal network of osimertinib intervention and highly activated *MNK1/2* were associated with L858R-positive lung adenocarcinoma. Upstream regulators and causal networks predicted suggested a close link to *EGFR* mutation-positive cancers, mainly NSCLC.

INTRODUCTION

The discovery of somatic mutations in the tyrosine kinase domain of the epidermal growth factor receptor (EGFR) (1, 2) drastically changed the therapeutic perspective of non-small-cell lung cancer (NSCLC). The representative *EGFR* oncogenic mutations are in-frame deletions in exon 19 (Ex19del) (44.8%) and a point mutation at Leu-858 substituted with arginine (L858R) (39.8%) (3). Personalized and/or precision medicine (PM) have been successful by targeting those mutations with tyrosine kinase inhibitors (TKIs) gefitinib, erlotinib, and afatinib. Because most patients, however, suffer from drug resistance after a year of treatment, therapeutic strategies have been challenged to improve the survival benefit of first-line treatment. The efficacy of the first- and second-generation EGFR-TKIs is limited by the result of drug resistance conferred by another mutation involving the substitution of threonine 790 with methionine (T790M) (4).

Osimertinib is an irreversible third-generation EGFR-TKI that is selective for sensitizing EGFR and T790M mutations. The randomized phase III AURA3 trial demonstrated that the efficacy of osimertinib was significantly greater than that of platinum therapy plus pemetrexed in patients with T790M-positive advanced NSCLC (5). Recently, osimertinib was recommended as first-line treatment for patients with *EGFR*-mutant NSCLC according to the FLAURA trial that reported significantly better PFS and OS with osimertinib than with first-generation EGFR-TKIs (gefitinib or erlotinib) (6, 7).

Numerous studies have been reported regarding *EGFR* mutations and their disease-related downstream signaling pathways (8–12) and EGFR-TKIs resistance (13–15). Hyperactivation of STAT3 enhances carcinogenesis in various cancers (16, 17) and drives drug resistance in response to EGFR TKIs (18). Chromosomal instability was found to be significantly increased during TKI treatment in T790M-negative patients and resulting co-acquired alterations and genomic evolution are primarily responsible for resistance to the first-generation TKIs (19). Low-frequency *EGFR* mutations in NSCLC, including point mutations, deletions, insertions, and duplications within exons 18–25, are also associated with poor responses to EGFR TKIs (20).

Today, the field of proteomics is strongly dominated by mass spectrometry (MS)–based methodologies, largely due to that modern mass spectrometers offer high mass resolution and accuracy required for correct protein identification. The most successful approach is "shotgun" proteomics that employs proteases (often trypsin) to enzymatically cleave proteins resulting in peptides, which are more convenient to separate and sequence (21). The most reliable protein identification strategy in shotgun proteomics is based on tandem (MS/MS) mass spectra generation of tryptic peptides by fragmentation and their consecutive search against databases of canonical/consensus sequences. MS-based proteomics has been extensively applied to investigate EGFR regulations, including phosphorylation, ubiquitination, and protein–protein interactions as well as post-translational modifications (22, 23). Zhang et al. performed quantitative phosphoproteomics to unveil global phosphorylation changes upon the erlotinib treatment of *EGFR* mutation-positive lung adenocarcinoma cells (24, 25).

Unfortunately, mutant proteins, those products of non-synonymous single nucleotide polymorphisms (nsSNP), are overlooked in general MS-based proteomic data because these proteoforms are excluded in canonical protein databases (26). Although, the high number of nsSNPs, estimated to be >3 million, suggests that single amino acid variants (SAAVs) are widely distributed in the human proteome (27), only a couple of mutant proteins have been detected at expression level in human samples (28). Cancerous diseases are often characterized by high mutation rates (29) that are tightly associated with the physiological and pathological traits of individuals (30), whereas the allele-specific gene expressions in the heterozygous state are also associated with various traits of individuals (31, 32).

Because many of these mutant proteins are exclusively expressed in cancer cells (33), they can be identified as lead candidates of optimal disease biomarkers. The qualitative and quantitative analyses of these proteoforms, thus, can provide novel diagnostic and prognostic values.

A laser microdissection (LMD) technique enables the collection of target cells of a certain type from sections of formalin-fixed paraffin-embedded (FFPE) cancer tissue (34, 35). Label-free spectral counting and identification-based semiquantitative shotgun proteomic analysis of microdissected target cancerous cells of a certain type were used that characterized lung adenocarcinoma (35).

A pivotal challenge is to understand how the major driver mutations—*EGFR* L858R and Ex19del—affect disease-related downstream networks together with other upstream driver mutation crosstalk, which plays a central role in the context of lung cancer progression, malignancy, and outcome and/or resistance regarding TKI therapies (28). We performed mutant proteomic analysis and applied the weighted gene correlation network analysis (WGCNA), which is an unsupervised gene-clustering method based on the correlation network of gene expression (36–38) as well as spectral counting-based comparative analysis. The main aim of this study was to identify

the key modules and networks of mutant proteins associated with the *EGFR* mutations L858R and Ex19del. To our knowledge, this is the first proteomics study performed to identify mutant proteoforms expressed in clinical tissue specimens.

MATERIALS AND METHODS

FFPE Tissue Specimens and Sample Preparation

Among 974 patients who underwent surgical lung cancer resection at St. Marianna University Hospital between 2000 and 2014, only 674 (69.3%) had tumors that were histologically confirmed adenocarcinomas. Pathological specimens were reviewed by pathologists to confirm that they satisfied the 2015 WHO classification of lung tumors (histological criteria) (39). For tissue microdissection, 10-μm-thick sections from the FFPE tumor blocks were cut onto DIRECTOR slides (OncoPlex Diagnostics Inc., Rockville, MD, USA). The sections were deparaffinized and stained only with hematoxylin using standard histological methods prior to dissection. Microdissection was performed using a Leica LMD7 Microdissection Microscope (Leica, Wetzlar, Germany). A total area of 4 mm^2 with about 15,000 tumor cells was transferred from the FFPE sections via laser dissection directly into the cap of a 200-μL low-binding tube. Proteins were extracted and digested with trypsin using Liquid Tissue MS Protein Prep kits (OncoPlex Diagnostics, Inc.). The procedures have been described in detail elsewhere (34, 35, 38).

Liquid Chromatography-Tandem Mass Spectrometry

Digested protein samples were used for liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis on a Q-Exactive Orbitrap mass spectrometer (Thermo-Fisher Scientific, Bremen, Germany) equipped with an LC system operated at 500 nL/min via a nano-ESI device (AMR Inc., Tokyo, Japan). The gradient was 110 min long and a 5-μL sample was injected in each analysis.

All LC-MS/MS data were acquired using Xcalibur, version 2.8 SP1 (Thermo Fisher Scientific) in high-resolution data-driven analysis (DDA) mode with the survey scan (MS in the mass range *m/z* 400–1,600) acquired in the Orbitrap at 70,000 resolution (at *m/z* 200) in profile mode. The survey scan was followed by the top 10 higher energy collision-induced dissociation (HCD) MS/MS spectra, acquired in centroid mode in the Orbitrap at 17,500 resolution.

For MS/MS acquisition of top 10 precursors, the following settings were used: minimal signal threshold = 1,700; isolation width = 1.6 *m/z*; normalized collision energy = 27%. Monoisotopic precursor selection, charge-state screening, and charge-state rejection were enabled with rejection of singly charged and unassigned charge states. Dynamic exclusion was enabled to remove selected precursor ions (±10 ppm) for 15 s after MS/MS acquisition. The expression levels of identified mutant proteins were assessed by spectral count-based protein quantification. Fold changes of expressed proteins in the base 2 logarithmic scale (R_{SC}) (40) were calculated using the spectral count (*SpC*) that is the number of MS/MS spectra assigned to each mutant protein.

Identification of Mutant Proteoforms

A strategy to identify mutant proteoforms in lung cancer samples was designed using high-quality shotgun proteomics tandem mass spectra. The central component of the approach was a unique set of protein sequences, which included SAAV sequences translated from known genomic studies. Using a custom-made software tool (FastaWriter v1.4.0), a new database of mutant protein sequences (ProteoFinder v17.04.12) was generated to include a mutation in each new entry that, thus, differed in amino acid from the consensus protein. Titin (Q8WZ42) was removed from the database to decrease the size of the database as 21,045 mutations were registered only on this protein. The resulting *in silico* derived proteoforms (total number of searchable mutations of 1,899,031) were denoted following the neXtProt nomenclature, including the access codes but adding information also about the position of the mutation (such as NX_P07288-S132L). These SAAV sequences were then shortened to reduce redundancy, keeping only the part of the protein sequence where the amino acid exchange took place surrounded by two additional tryptic peptides at both N- and C-termini. The new database entries were rendered as a combination of consensus (neXtProt database 2017-04-12 release) and mutant proteoform sequences in standardized *fasta* format. **Figure 1** illustrates a general workflow of identification of SAAVs by tandem mass spectra searching a specialized protein database, ProteFinder (PFdb), and MS-based sequencing of a mutant peptide is exemplified.

The MS raw files of 108 runs (36 patient samples as triplicate) were imported into PEAKS Studio v8.5 (Bioinformatics Solutions Inc., Waterloo, Canada) (41) for database searching against the PF v17.04.12 database, appended with contaminant sequences (cRAP). PEAKS database searches were performed with a precursor ion error tolerance of 10 ppm, fragment ion error tolerance of 0.05 Da, fixed carbamidomethyl cysteine, and modifications of oxidation (M), deamidation (NQ), and acetylation (N-term) were set dynamically. Trypsin was specified as the enzyme, allowing for two missed cleavages. The technical triplicates were searched together, generating a single combined result file of each biological sample.

The search results were further filtered for hits with mutant specific tryptic peptides removing all multiple protein identifications while multi-isoform hits with the same amino acid change were included in the final list. Non-tryptic peptides with the mutation were not considered as reliable identification and were excluded in the additional filtering steps. Isobaric amino acid mutations, i.e., exchange of Ile to Leu and Leu to Ile, were registered for future experimental verification and kept as potentially valid identifications. The summary of each search, including score distributions and statistical data, which are available as PDF files (e.g., AZ0x_summary.pdf) in Search summaries in **Supplementary Information File 1**.

The technical triplicates together have resulted in rich data with an MS/MS spectra range of 44,798–143,118, providing

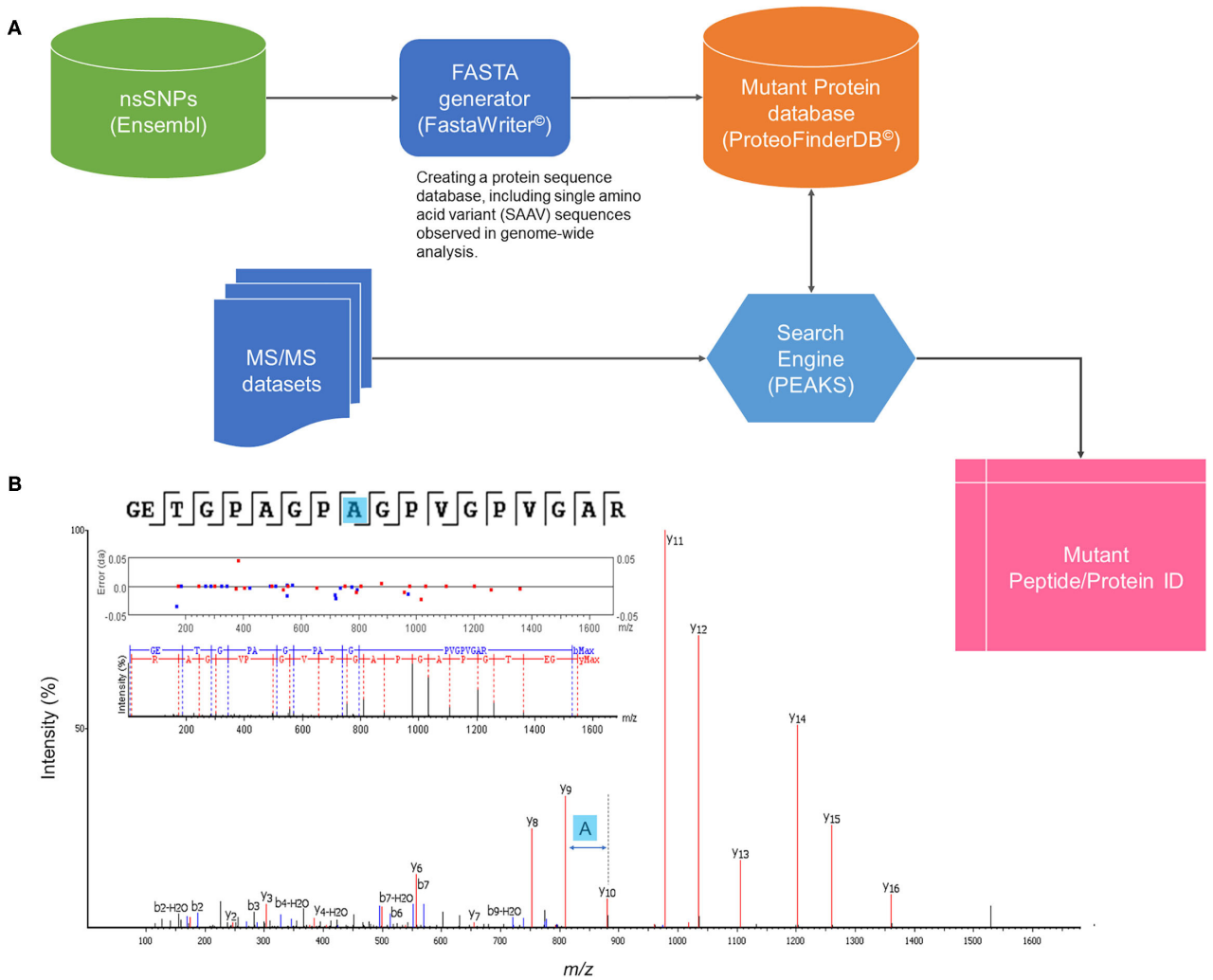


FIGURE 1 | General workflow of mutant proteomics. **(A)** Identification of SAAVs by tandem mass (MS/MS) spectra searching a specialized protein database, ProteFinder (PFdb), which includes 1.9 million SAAVs, all human consensus sequences, splicing variants, and 105 common contaminant proteins. PFdb is verified in PEAKS (Biosystems Inc.) and used together with its decoy databases. **(B)** MS-based sequencing of a SAAV of collagen a-1(I) chain (COL1A1–P02452) at the point mutation, T1075A.

a peptide sequence match (PSM) range of 10,464–51,133, peptide sequence range of 7,703–19,974, and protein group range of 1,221–2,266. The identified mutant protein sequences were between 252 and 964 after filtration, which is presented in **Supplementary Information File 2**. The protein sequences carrying amino acid variants were registered, and the presence of each mutant protein was indicated in the technical replicates as well as their scan numbers.

WGCNA

The similarity among protein expression patterns for all protein pairs was calculated according to their pairwise Pearson's correlation coefficient, i.e., the similarity between proteins i and j was defined as $(1 - r_{i,j})/2$, where $r_{i,j}$ is the correlation of the protein expression patterns between the two proteins i and j. An adjacency matrix was then computed by increasing the similarity matrix up to the power of 10 to generate a coexpression network with scale-free properties. Subsequently, from the resultant scale-free coexpression network, we generated a topological overlap matrix (TOM) that considers topological similarities between a pair of proteins in the network. Using the dissimilarity according to TOM (1 – TOM), we conducted hierarchical clustering to generate a tree that clustered proteins in its branches. Dynamic tree cutting was used to trim the branches to identify protein modules. A protein module was summarized by the top hub protein (referred to as eigen-protein) with the highest connectivity in the module. To identify the protein modules associated with clinical traits, we calculated the correlation coefficients between the eigen-proteins and clinical traits. WGCNA was conducted using a Garuda Platform gadget (The Systems Biology Institute, Tokyo, Japan) that implemented the WGCNA pipeline based on the WGCNA R-package (36).

Protein–Protein Interaction Network Construction and Functional Enrichment

To construct a protein interaction network for a protein module, we used the STRING database (version 11.0) (42), which accumulates information on protein–protein interactions from various other databases, such as IntAct, Reactome, DIP, BioGRID, MINT, KEGG, NCI/Nature PID, the Interactive Fly, and BioCyc. STRING networks were constructed under the criteria for linkage only with experiments, databases, text mining, and coexpression using the default settings, i.e., a medium confidence score of 0.400, a network depth of 0 or 50 interactions. Subsequently, protein networks imported from the STRING database were visualized using *Cytoscape* version 3.7.2. Functional enrichment results were obtained for canonical pathways considering $p < 0.05$ to be statistically significant.

Comparative Analysis of the Causal Networks and Pathways Predicted by IPA

Canonical pathways, upstream regulators, and causal networks were predicted using the ingenuity pathway analysis (IPA) software (43). Mutant protein expression data (quantile-normalized for selected modules) were used as input data sets. Comparative analysis of the predicted causal networks (p-value < 0.05) was performed to elucidate networks associated with the three clinical traits: Ex19del, L858R, and no Ex19del/L858R mutations, where activation and inhibition of a predicted network were defined by z-values >2.0 and < -2.0, respectively, and upregulation and downregulation were defined by z-values >1.0 and < -1.0, respectively.

RESULTS

Mutant Proteome Data Sets of Lung Adenocarcinoma

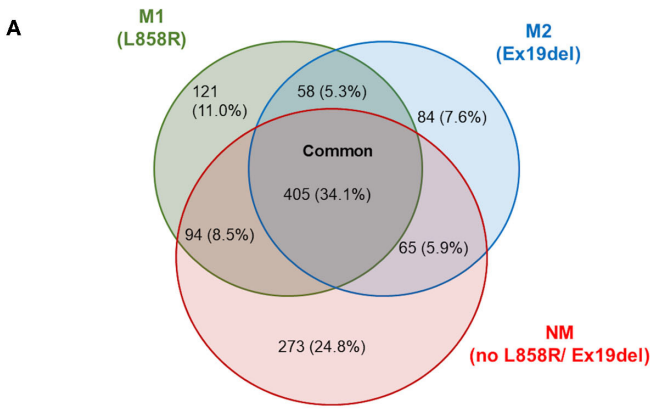
MS-based proteomic analysis was conducted for 36 FFPE tissue specimens of lung adenocarcinoma (35 involved the acinar subtype and one involved the papillary subtype). These specimens were selected for their preserved condition, tumor area, and well-clarified pathological diagnosis and *EGFR* mutation status (nine specimens of the clinical trait M1: L858R mutation, nine specimens of the clinical trait M2: Ex19del mutation, and 18 specimens of the clinical trait NM: no Ex19del or L858R mutation; see **Table 1**). Pre-surgical treatment was not performed in any of the cases.

A total of 1,100 mutant proteins were identified, in which M1, M2, and NM were 678, 612, and 837, respectively, and 405 (34.1%) were expressed commonly (**Figure 2A**). The proportion of mutant proteins unique to the L858R mutation was 121 (11.0%), and that to the Ex19del mutation was 84 (7.8%), whereas the proportion of proteins expressed in only no *EGFR* mutation cases was 273 (24.8%). GO analysis using PANTHER Ver. 14.1 (44) exhibited mostly similar profiles in gene hits for all the traits (M1, L858R mutation; M2, Ex19Del mutation; NM, no Ex19del or L858R mutation; see **Figure S1**). Mutation proteins with high hits in GO biological process included cellular process (GO:0009987), localization (GO:0051179), cellular component organization or biogenesis (GO:0071840), biological regulation (GO:0065007), metabolic process (GO:0008152), and response to stimulus (GO:0050896). Those in the GO protein class included cytoskeletal protein (PC00085), nucleic acid-binding protein (PC00171), and metabolite interconversion enzyme (PC00262).

An orthogonal partial least square-discriminant analysis (OPLS-DA) (45) was applied to identified mutant proteins and interestingly exhibited profound differences in distance among

TABLE 1 | Clinicopathological information of the 36 patients.

Variable	Category	No. patients	%
Gender			
	Female	16	44.4
	Male	20	55.6
AGE (years)			
	Median (range)	67.9 (22–82)	
Smoking index (Brinkmann Index, BI)			
	Female		
	BI = 0	12	75
	0 < BI ≤ 400	0	0
	400 < BI ≤ 600	0	0
	600 < BI ≤ 1,200	4	25
	BI > 1,200	0	0
	Male		
	BI = 0	3	15
	0 < BI ≤ 400	4	20
	400 < BI ≤ 600	0	0
	600 < BI ≤ 1,200	10	50
	BI > 1,200	3	15
Histologic type			
	Adenocarcinoma	36	100
Subtype			
	Acinar	35	97.2
	Papillary	1	2.8
Surgical method			
	Radical lobectomy	24	66.7
	Limited resection	12	33.3
Tumor size on CT			
	T1a (≤1 cm)	1	2.8
	T1b (1–2 cm)	11	30.6
	T1c (2–3 cm)	11	30.6
	T2a (3–4 cm)	4	11.1
	T2b (4–5 cm)	7	19.4
	T3 (5–7 cm)	2	5.6
	T4 (>7 cm)	0	0
Clinical stage			
	cIA	21	58.3
	cIIA	2	5.6
	cIB	9	25
	cIIB	1	2.8
	cIV	3	8.3
***EGFR* mutation status**			
	Positive		
	L858R	9	25
	Ex19 E746-A750 del	9	25
	Negative		
	Neither L858R nor Ex19del	18	50



Clinical trait	number of proteins
M1 (L858R)	678
M2 (Ex19del)	612
NM (Neither L858R nor Ex19del)	837
Overall number of proteins identified	1100

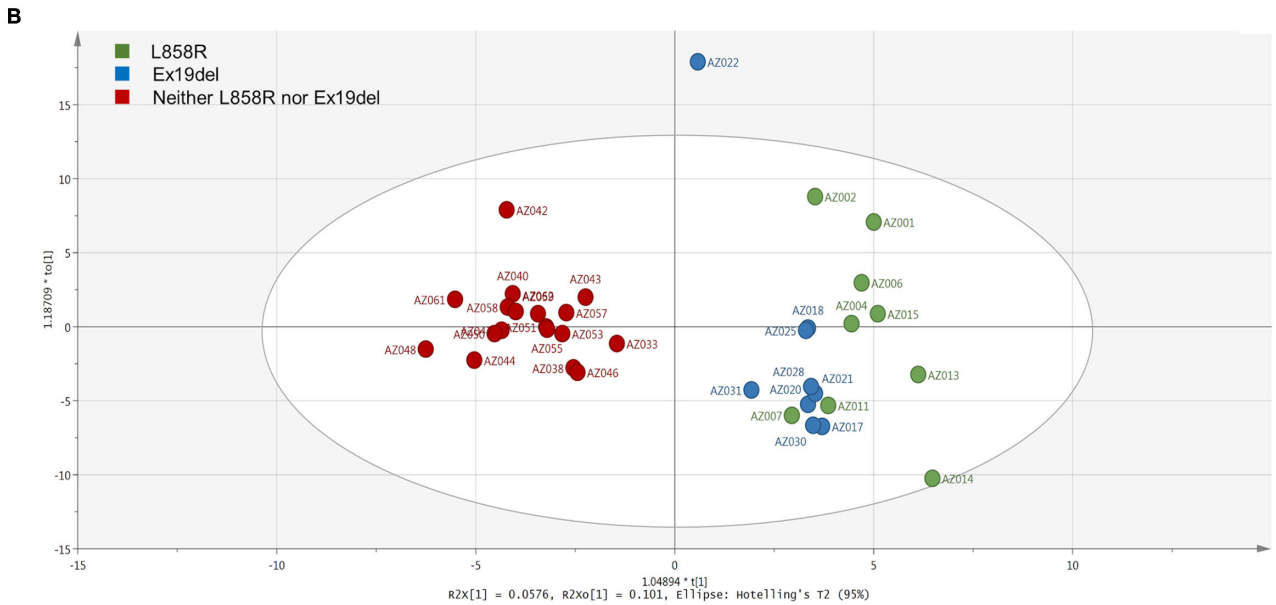


FIGURE 2 | Venn map and orthogonal partial least square-discriminant analysis (OPLS-DA) of the identified proteins. **(A)** Venn map of the identified proteins. **(B)** OPLS-DA of the expressed proteins including their spectral counts for patients.

the *EGFR* mutation statuses (**Figure 2B**), whereas a conventional hierarchical clustering of patients according to mutant protein abundance failed to show a clear separation among the three clinical traits. Surprisingly, clear differentiation was found between the NM group and the M1 and M2 groups. The data points of the M1 group appeared to be to some extent scattered, whereas those of the M2 group clustered closely. These findings seem to unveil the mutant proteome landscape correlating with the *EGFR* mutation type in lung adenocarcinoma.

Identification of Key Mutant Protein Modules by WGCNA

A weighted gene coexpression network was constructed in which all the identified mutant proteins were clustered, and we found 23 mutant protein modules (**Figures 3A,B**). A spectral counting-based heat map (46) for eigen-proteins in the modules is shown in **Figure 3C**. In the WGCNA, a soft threshold power of 15 was selected to define the adjacency matrix according to the criteria of approximate scale-free topology with a minimum module size of 30 and a module detection sensitivity *deepSplit* of 4. The clinical traits for patients were set according to the *EGFR* mutation status with M1, M2, and NM traits corresponding to L858R mutation, Ex19del mutation, and neither Ex19del/L858R mutation, respectively. The correlations between resultant modules and clinical traits were determined to identify mutant protein modules whose expressions were up- or downregulated in L858del, Ex19del, or no Ex19del/L858R mutation samples (**Figure S2**).

Among the 23 modules, only the WM6 module was moderately correlated with the *EGFR* Ex19del mutation status ($r = 0.41$, $p < 0.05$). Most of the other WGCNA modules were not statistically significant. However, several modules seem to be characteristic to the clinical traits (**Figure S2**). The WM10, WM12, WM14, and WM22 modules seem to be characteristic

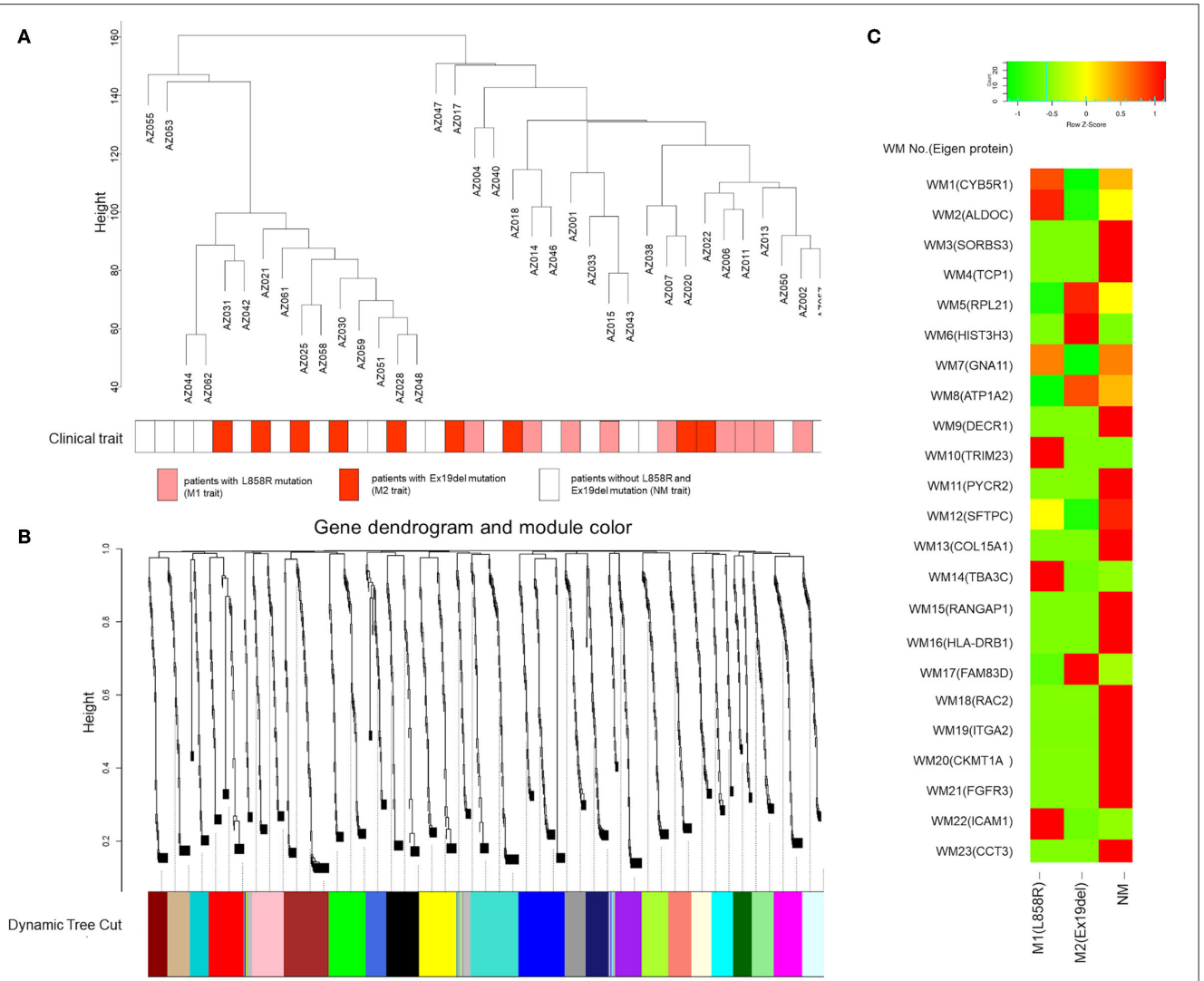


FIGURE 3 | Gene modules identified by weighted gene coexpression network analysis (WGCNA). **(A)** Patient clustering according to mutant protein abundance with the *EGFR* mutation profiles. The red, orange, and white cells below the patients indicate the *EGFR* mutation types, i.e., Ex19del mutation, L858R mutation, and no *EGFR* mutation, respectively. **(B)** Gene dendrogram obtained by clustering dissimilarity according to topological overlap with the corresponding module. The colored rows correspond with the 23 modules identified by dissimilarity according to topological overlap. **(C)** Heat map for the proteome abundance of eigen proteins in the 23 mutant protein modules by WGCNA. The rows and columns are the mutant protein modules and *EGFR* mutation types, respectively. The red and green colors indicate high and low mutant protein abundances, respectively, of an eigen protein in a mutant protein module. The names of the eigen proteins in the protein modules are indicated in parentheses.

to the L858R mutation status ($r = 0.3$, $p = 0.08$). The WM17 module showed a positive correlation with the Ex19del mutation status ($r = 0.3$, $p = 0.08$). We could find no modules characteristic to the NM trait (no L858R or Ex19del mutations).

WGCNA Modules Screened by ORA and Functional Enrichment Analysis

The computational WGCNA framework (36) has been proven to be powerful in identifying coexpression protein modules (37, 38). However, it should be noted that traditional trait analysis of the correlations between eigen components of WGCNA modules and clinical traits might overlook important modules for investigating molecular mechanisms differentially behind a disease. Especially for clinical traits quite close to each other, difficulties would be sometimes encountered to attain identification of key WGCNA modules with a high significance. Multiple correction testing, such as Bonferroni, Benjamini-Hochberg, etc., would result in that none of the modules associated with M1 or M2 remains significant. Statistical over-representative analysis (ORA) would help to evaluate potential key WGCNA modules with identified proteins uniquely expressed and upregulated to each trait.

We conducted an ORA-based screening of WGCNA mutant protein modules to further identify key protein modules to

A

Group	L858R mutation	Ex19del mutation	No L858R/Ex19del mutation	# of proteins	*Notes*:	Modules
Group 1	Uniquely identified	N/A	N/A	121	Characteristically expressed under L858R mutation (M1)	WM10, WM14, WM22,
Group 2	**Up-regulated***	N/A	N/A	132		
Group 3	N/A	Uniquely identified	N/A	84	Characteristically expressed underEx19del mutation (M2)	WM6, WM17
Group 4	N/A	**Up-regulated***	N/A	142		
Group 5	N/A	N/A	Uniquely identified	273	Characteristically expressed under no L858R/Ex19del mutations (NM)	WM4, WM9, WM11, WM20, WM21, WM23

*protein group upregulated with $|R_{SC}| > 1$ (M1>M2 or M1< M2)

B

Module ID	Module color	Group 1 (*q*-value)	Group 2 (*q*-value)	Group 3 (*q*-value)	Group 4 (*q*-value)	Group 5 (*q*-value)	Max *q*-value
WM14	Midnightblue	0.00277242	8.373E-09	1.00000	0.03824	1.00000	8.37331E-09
WM22	Brown	2.9532E-08	1.841E-07	1.00000	0.52180	0.22080	2.9532E-08
WM10	Turquoise	1.451E-05	0.11910	0.03833	0.41549	1.00000	1.451E-05
WM6	Blue	1.00000	0.00291	2.661E-05	1.560E-04	1.00000	2.661E-05
WM17	Lightgreen	1.00000	0.03262	0.00303	5.66507E-05	1.00000	5.665E-05
WM9	Tan	1.00000	0.01991	1.00000	0.13625	1.522E-04	1.522E-04
WM4	Red	0.06133	0.21735	1.00000	0.29685	1.594E-04	1.522E-04
WM23	Darkred	1.00000	0.29316	1.00000	0.26736	0.00038	0.00038
WM11	Salmon	0.10120	0.09636	1.00000	0.28620	0.00066	0.00066
WM21	Darkturquoise	1.00000	0.14919	1.00000	0.42164	0.00069	0.00069
WM3	Black	0.00226	0.42766	0.00157	0.15398	1.00000	0.00157
WM20	Greenyellow	1.00000	0.13897	1.00000	0.22695	0.00329	0.00329
WM13	Purple	0.97018	0.12423	1.00000	0.10464	0.00343	0.00343
WM2	Yellow	0.00767	0.09852	0.01150	0.04682	1.00000	0.00767
WM8	Green	1.00000	0.01264	0.00920	0.09362	0.38717	0.00920
WM15	Magenta	1.00000	0.04400	1.00000	0.09362	0.03163	0.03163
WM12	Royalblue	0.04271	0.44590	1.00000	0.03553	1.00000	0.03553
WM5	cyan	1.00000	0.08086	0.24314	0.03960	3.01300	0.03960
WM7	pink	0.05463	0.27382	1.00000	0.06169	0.70479	0.05463
WM16	lightyellow	1.00000	0.46819	1.00000	0.52809	0.23767	0.23767
WM1	darkgreen	0.89317	0.26700	1.00000	0.11877	0.53785	0.26700
WM18	lightcyan	1.00000	0.54552	1.00000	0.31698	0.28227	0.28227
WM19	grey60	1.00000	0.52180	1.00000	0.09769	0.76207	0.52180

FIGURE 4 | Overlapping proteins unique to the clinical traits and/or upregulated under the M1 or M2 traits and those from weighted gene coexpression network analysis (WGCNA). **(A)** Results of identified proteins and spectral-counting based semiquantitative comparison. Each row represents results for each protein group. The red and pink cells in the "L858R mutation" and "Ex19del mutation" columns indicate that the proteins in the group are uniquely expressed and significantly upregulated, respectively, in samples with the mutations [Upregulated with $|Rsc| > 1$ (M1 > M2 or M1 < M2)]. The fourth column shows the number of proteins in each protein group. The fifth column provides notes for each protein group. The WGCNA modules with significant overlap with each protein group are listed in the sixth column ("Modules" column). **(B)** Overlap in proteins between the groups by the protein expression profiles and the modules by WGCNA. Each row in the embedded table represents overlap analysis results for each module. The first and second columns in the table represent module ID and color name of the module. The third through eighth columns indicate the *q*-values for overlap in proteins between a module by WGCNA and the five protein groups. In the six columns, significant *q*-values are highlighted in red. The eighth column represents the value of the most significant *q*-value (max *q*-value) in each module. The 18 modules with max *q*-values <0.05 are listed in order.

investigate the differential disease mechanisms associated with the *EGFR* L858R and Ex19del mutation statuses; 121, 84, and 273 mutant proteins identified were expressed uniquely to the respective traits: M1, M2, and NM (**Figure 3A**); 132 and 142 mutant proteins were upregulated differentially to M1 and M2 with $|R_{SC}| > 1$ (higher than 2-fold change) in the comparison between M1 and M2 (**Figure S3**). The overlaps between the WGCNA-derived protein modules and identification-based significantly expressed proteins were then assessed using the over-representation test. We confirmed that five WGCNA modules overlapped significantly (maximum *q*-value among the groups <0.05) with protein groups unique to each trait and/or highly upregulated to M1 or M2 (**Figures 4A,B**).

To characterize those five modules, we analyzed the biological connectivity among the proteins in each module by mapping the module proteins in the human protein–protein interaction (PPI) network and among the biological pathways by pathway enrichment analysis (**Figures 5**, **6**).

Three WGCNA modules—WM10, WM14, and WM22—significantly overlapped with protein groups uniquely identified and highly upregulated under the L858R mutation (**Figure 4**). The enriched pathways of the WM10 mutant protein module involved DNA duplex unwinding, canonical glycolysis, glucose catabolic process to pyruvate, DNA unwinding involved in DNA replication, COPI-dependent Golgi-to-ER retrograde traffic, the role of GTSE1 in G2/M progression after G2 checkpoint,

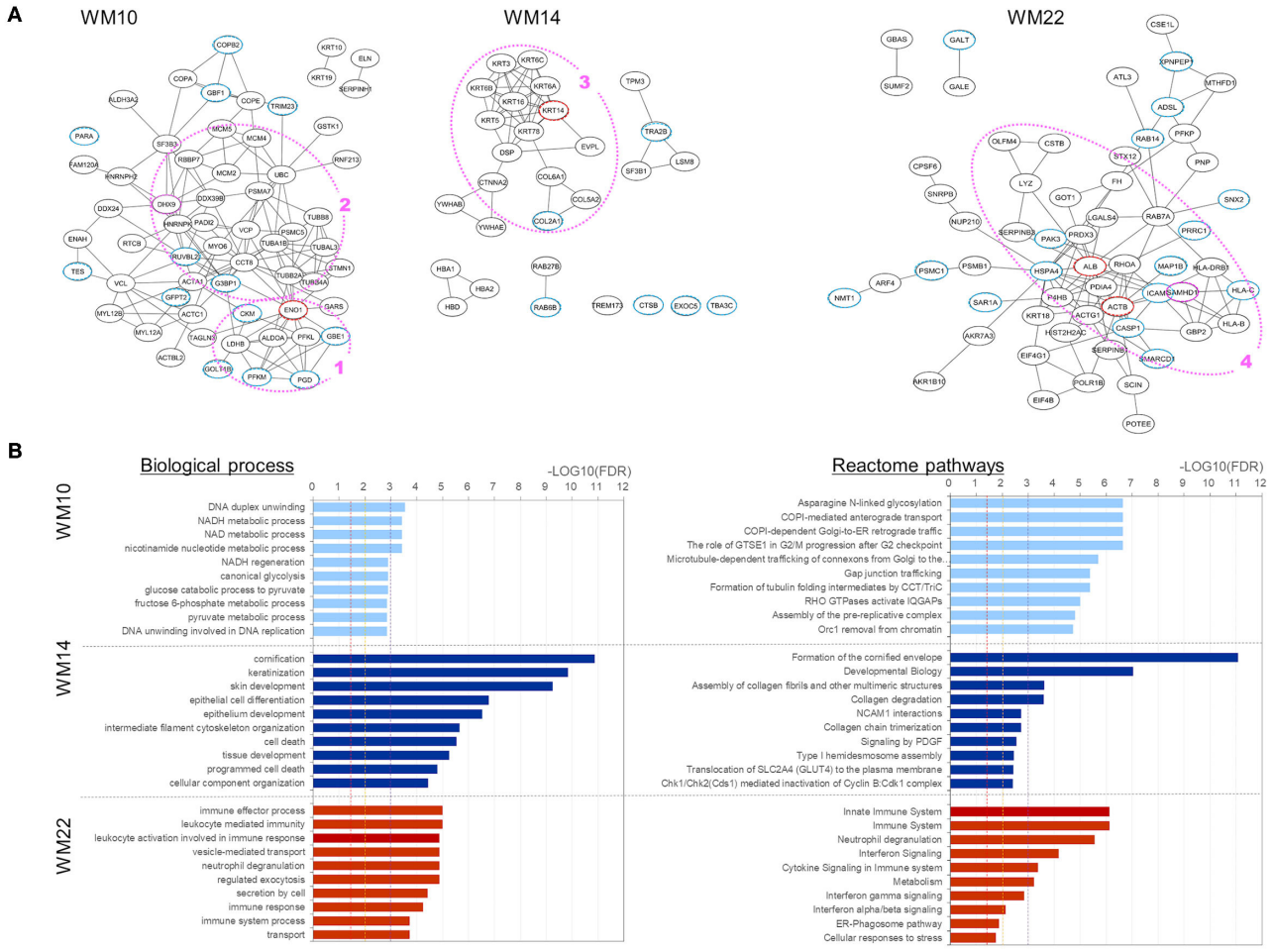


FIGURE 5 | Analysis results for three protein modules (WM10, WM14, and WM22) that overlap with proteins uniquely and upregulated under the L858R mutation, respectively. **(A)** Protein interaction networks for the three WGCNA modules. Dotted circle nodes in blue and red represent eigen-proteins and hub proteins, respectively, for each module. **(B)** Pathway enrichment analysis using Go Biological Process and Reactome pathway databases for the three protein modules. The vertical axis shows the pathway names, and the bars on the horizontal axis represent the –log10 (*p*-value) of the corresponding pathway. The different colors of the bars are following the corresponding modules. Dashed lines in red, orange, and magenta indicate *p*-values <0.05, <0.01, and <0.001, respectively.

and formation of tubulin folding intermediates by CCT/TriC (**Figure 5B**). The hub protein alpha-enolase (also known as MBP-1) encoded by ENO1 is involved in the subnetworks related to the carboxylic acid metabolic process (as indicated by the pink dotted line 1 in **Figure 5A**) and is associated with glycolysis, growth control, and hypoxia tolerance. MBP-1 binds to the myc promoter and acts as a transcriptional repressor and so maybe a tumor suppressor. The cell cycle–related subnetwork is denoted by the pink dotted line 2 in **Figure 5A**. DHX9 encodes ATP-dependent RNA helicase A [also known as nuclear DNA helicase II (NDH II) or leukophysin (LKP)] participates in multiple processes of gene regulation, including transcription, translation, and DNA replication, and plays important roles at the maintenance of genomic stability. DHX9 has been reported to be overexpressed in various types of malignant tumors and might be a potential therapeutic target for the treatment of NSCLC (47).

The enriched pathways of the WM14 module include epithelial cell differentiation, tissue development, cell death, programmed cell death, developmental biology, and collagen degradation (**Figure 5**). The hub protein KRT14 is associated with developmental biology, which subnetwork is indicated by the pink dotted line 3 in **Figure 5A**. The enriched pathways of the WM22 module involve the immune effector process, immune response, cytokine signaling in immune system, and cellular responses to stress (**Figure 5B**). The subnetworks related mostly to the immune system (the pink dotted line 4 in **Figure 5A**), in which sterile alpha motif and HD domain-containing protein 1 (SAMHD1), a deoxyribonucleoside triphosphate triphosphohydrolase is known to play roles in defense response to the virus and cellular response to DNA damage stimulus, and is dysregulated in breast and other cancers (48). Frequently mutated SAMHD1 found in colon cancers was suggested to be involved in tumorigenesis with defective mismatch repair (MMR) (49) and also act as a resistance factor for anticancer drugs (50).

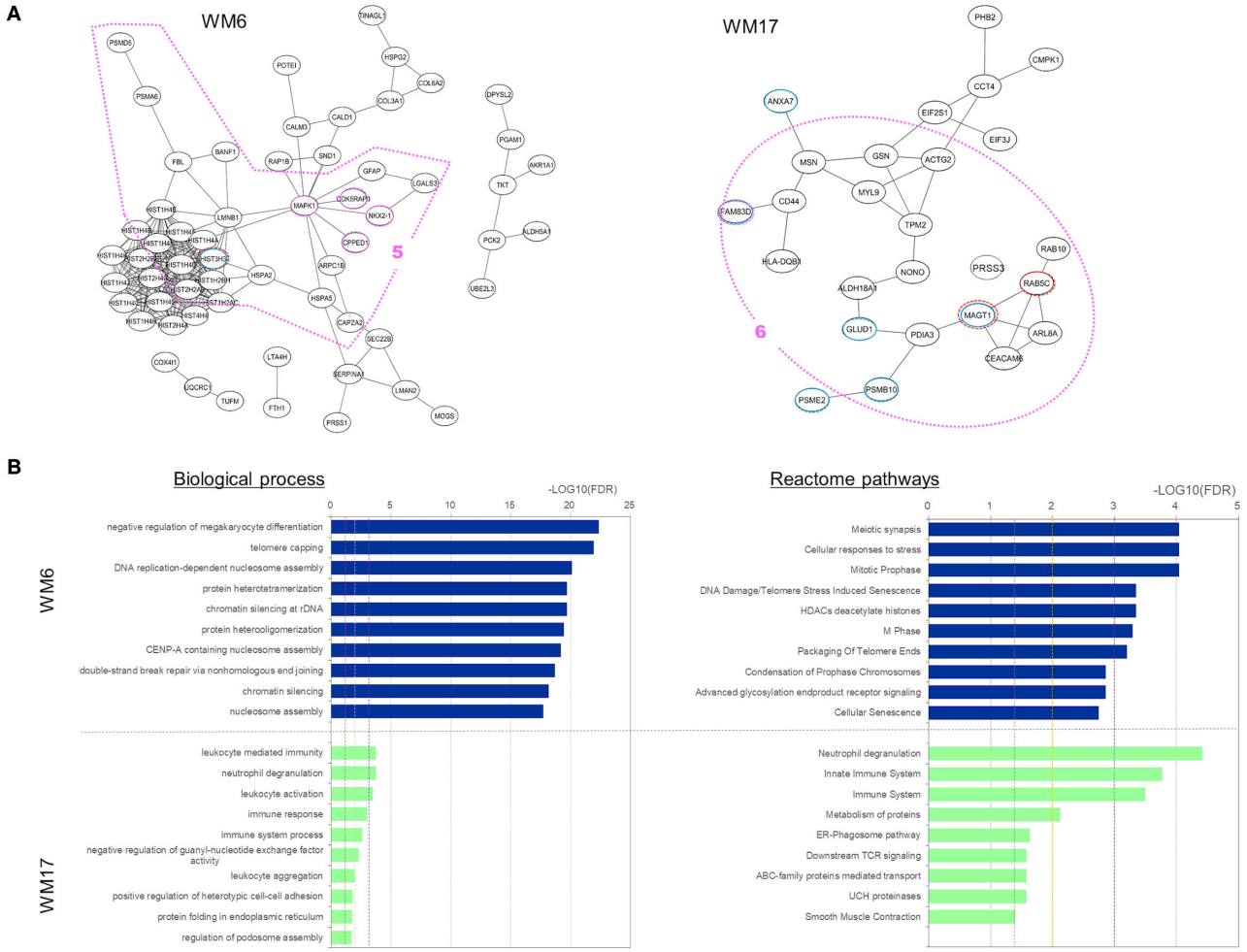


FIGURE 6 | Analysis results for three protein modules (WM6 and WM17) that overlap with proteins uniquely expressed and upregulated under the Ex19del mutation. **(A)** Protein interaction networks for the three WGCNA modules. Dotted circle nodes in blue and red represent eigen-proteins and hub proteins, respectively, for each module. **(B)** Pathway enrichment analysis using Go Biological Process and Reactome pathway databases for the three protein modules. The vertical axis shows the pathway names, and the bars on the horizontal axis represent the –log10 (*p*-value) of the corresponding pathway. The different colors of the bars are following the corresponding modules. Dashed lines in red, orange, and magenta indicate $p < 0.05$, <0.01, and <0.001, respectively.

Two WGCNA modules—WM6 and WM17—significantly overlapped with protein groups uniquely identified and highly upregulated under the Ex19del mutation (**Figure 4**). The enriched pathways of the WM6 module involved DNA replication-dependent nucleosome assembly, chromatin silencing, double-strand break repair via non-homologous end joining, cellular responses to stress, DNA damage/telomere stress-induced senescence, and M phase (**Figure 6B**). The hub protein is the mutant H3.1t encoded by the mutant HIST3H3. Histone H3.1t protein (also known as H3t) itself is a core component of the nucleosome and plays a central role in transcription regulation, DNA repair, DNA replication, and chromosomal stability. The subnetworks related to both cellular responses to stress and mitotic prophase are indicated by the pink dotted line 5 in **Figure 6A**. Calcineurin-like phosphoesterase domain containing 1 (CPPED1, also known as CSTP1) blocks cell cycle progression and promoting cell apoptosis by dephosphorylating AKT family kinase (51). CDK5RAP3 itself encodes CDK5 regulatory subunit associated protein C53 (Cdk5rap3, also known as C53 and LZAP) that is a probable tumor suppressor involved in signaling pathways governing transcriptional regulation and cell cycle progression. Its specific mutant protein was reported to prevent apoptosis-induced cleavage of nuclear substrates, including nuclear shrinkage, chromatin condensation, and DNA fragmentation (52). The homeobox protein Nkx-2.1 [also known as thyroid transcription factor-1 (TTF-1)] has a role in lung development and surfactant homeostasis and is highly expressed in both small-cell lung carcinoma (SCLC) and lung adenocarcinoma (53, 54). Based on a quantitative real-time RT-PCR study of the NSCLC cell lines, Zu et al. (55) concluded that TTF-1 may serve as a tumor suppressor because of its inverse correlation with Ki-67 proliferative activity and increase of cellular apoptosis.

TABLE 2 | The top master regulators of causal networks predicted using the ingenuity pathway analysis (IPA) for the WM10 and WM6 modules, which are representatively characteristic under the L858R and Ex19del mutation statuses, respectively.

Module ID (module color)	Master regulator	Molecule type	Depth	Predicted activation state	Activation *z*-score	*p*-value of overlap	Network bias-corrected *p*-value	Participating regulators[a]	Target molecules in dataset[b]
WM 10 (turquoise)	Osimertinib	Chemical drug	3	Downregulated highly under M1 and NM	−1.667	3.84E-12	1.00E-04	*AGT, Akt, AKT1, BLK, Creb, CTNNB1, EGFR, ERBB2, ERBB3, ERBB4, ERK1/2, ESR1, FOS, FOXO1, HBEGF, HIF1A, HRAS, JUN, KRAS, MAP2K1, MTOR, NFkB (complex), NRG1, osimertinib, PSEN1, RAF1, RB1, ICTOR, SRC, STAT1, STAT3, STA T5B, STAT6, TNK2, TP53, TSC2*	*ACADVL, ACTA1, ACTC1, ALDH3A2, ALDOA, COPA, COPB2, COPE, DDX39B, ENAH, ENO1, FAM120A, G3BP1, GBE1, GFPT2, GSTK1, HNRNPK, KRT10, LDHB, MACROH2A1, MCM2, MCM4, MCM5, MMUT, NIPSNAP1, PADI2, PDXDC1, PFKL, PFKM, PGD, PRPH, PSMC5, PURA, RBBP7, RNF213, RTCB, RUVBL2, SERPINH1, SF3B3, STMN1, TES, TMEM109, TUBB2A, TUBB4A, VCL, VCP*
	NEUROG1	Transcription regulator	3	Downregulated highly under M1 and NM	−1.89	4.35E-11	1.00E-04	*Akt, AKT1, Cdc42, CHUK, ERK1/2, HRAS, MAP2K1/2, MRTFA, NEUROG1, PI3K(complex), RAF1, RELA, RHOA, ROCK2, SRF, STAT, STAT3, STAT4, STAT5A, STAT5B, STAT6, TP53*	*ACADVL, ACTA1, ACTC1, ALDH3A2, ALDOA, COPB2, ELN, ENAH, ENO1, FAM120A, GBE1, GFPT2, GSTK1, KRT10, LDHB, MACROH2A1, MCM2, MCM4, MCM5, MMUT, MYO6, NOL3, PADI2, PFKM, PGD, PURA, RBBP7, RTCB, RUVBL2, SERPINH1, STMN1, TES, TUBB2A, UBC, VCP*
	PARPBP	Other	3	Activated highly under M1	2.449	4.52E-11	1.00E-04	*AKT1, ESR1, ESR2, estrogen receptor, NFkB (complex), PARPBP, PI3K (complex), PRL, STAT1, STAT4, STAT5A, STAT5B, TP53*	*ACADVL, ACTC1, ALDH3A2, ALDOA, CKM, COPA, COPE, DDX39B, ELN, ENAH, FAM120A, GBE1, GFPT2, HNRNPH2, KRT19, MACROH2A1, MCM2, MCM4, MCM5, MMUT, NOL3, PADI2, PFKM, PGD, PURA, RBBP7, RNF213, RUVBL2, S100A7A, SERPINH1, STMN1, TUBA1B, UBC, VCL, VCP*
	HOXA1	Transcription regulator	3	Activated highly under M1	2.646	5.38E-11	1.00E-04	*AKT1, EGFR, ESR1, ESR2, HOXA1, PI3K (complex), PPARA, PRL, RELA, STAT1, STAT4, STAT5A, STAT5B, TP53*	*ACADVL, ACTA1, ACTC1, ALDH3A2, ALDOA, CKM, COPA, COPE, DDX39B, ELN, ENAH, FAM120A, GBE1, GSTK1, HNRNPH2, KRT19, MACROH2A1, MCM2, MCM4, MCM5, MMUT, NOL3, NQO1, PADI2, PFKM, PGD, PURA, RNF213, RUVBL2, S100A7A, SERPINH1, STMN1, TUBA1B, UBC, VCL, VCP*
	AMOT	Other	3	Inhibited highly under M1 and M2	−2.646	7.02E-11	1.00E-04	*AKT1, AMOT, CDKN1A, CTNNB1, DYRK1A, ERBB4, GLI1, ITCH, Jnk, JUN, JUNB, LATS2, MAP3K7, NFkB (complex), RAF1, RELA, TAZ, TP53, TP63, TP73, YAP1*	*ACADVL, ACTA1, ACTC1, ALDOA, CNIH4, FAM120A, GBE1, GFPT2, HLA-, KRT10, KRT19, LDHB, MACROH2A1, MCM2, MCM4, MCM5, MMUT, MYL12A, MYO6, NOL3, NQO1, PADI2, PFKM, PGD, PURA, RBBP7, RTCB, SERPINH1, SF3B3, TUBB2A, UBC, VCL, VCP*
	IP6K2	Kinase	2	Inhibited highly under M1 and M2	−2.236	8.93E-11	1.00E-04	*Akt, AKT1, IP6K2, MAP3K7, STK11, TP53*	*ACADVL, ACTA1, ALDOA, CKM, ENO1, FAM120A, GSTK1, HLA-B, LDHB, MACROH2A1, MCM2, MCM4, MCM5, MMUT, MYO6, NOL3, NQO1, PADI2, PFKM, PGD, PURA, RBBP7, SERPINH1, STMN1, UBC, VCL, VCP*

(Continued)

TABLE 2 | Continued

Module ID (module color)	Master regulator	Molecule type	Depth	Predicted activation state	Activation *z*-score	*p*-value of overlap	Network bias-corrected *p*-value	Participating regulators[a]	Target molecules in dataset[b]
	APH-1	Group	3	Activated highly under M1	2	9.86E-11	1.00E-04	*Akt, AKT1, APH-1, APH1A, APH1B, NOTCH1, PSEN1, PSEN2, Secretase gamma, TP53*	*ACADVL, ALDOA, CKM, ENO1, FAM120A, GSTK1, HLA-B, LDHB, MACROH2A1, MCM2, MCM4, MCM5, MMUT, MYO6, NOL3, NQO1, PADI2, PDXDC1, PFKM, PGD, PURA, RBBP7, SERPINH1, STMN1, UBC, VCL, VCP*
WM6 (blue)	*ASGR1*	Transmembrane receptor	3	Activated highly to M2	3.606	1.90E-11	1.00E-04	*AGT, AKT1, ASGR1, CCND1, CEBPB, CREB1, CTNNB1, DDIT3, EGFR, ERBB2, ERK, ERK1/2, ESR1, estrogen receptor, FOS, FOXO1, HBEGF, HIF1A, HRAS, JUN, MKNK1, MTOR, MYC, MYCN, NCOA2, NOS2, PGR, PLCB1, PPARG, PTPN11, SIRT1, SRF, TP53*	*AKR1A1, ARPC1B, ATP6V0D1, CAPZA2, CDK5RAP3, CLDN3, COL3A1, COL6A2, DPYSL2, FBL, GPD2, H2AC6, H2BC18, H4C1, H4C11, H4C12, H4C14, H4C15, H4C2, H4C3, H4C4, H4C6, H4C8, H4C9, HSPA2, HSPA5, HSPG2, LMAN2, MAPK1, MOGS, NKX2-1, PCK2, PGAM1, PRSS1, PSMA6, PTGFRN, RBM3, SEC22B, SELENBP1, SND1, TINAGL1, TKT, UBE2L3, UQCRC1*
	CEBPB	Transcription regulator	1	Activated commonly	3.317	1.78E-10	1.00E-04	*CEBPB*	*GFAP, H4C1, H4C11, H4C12, H4C14, H4C15, H4C2, H4C4, H4C6, H4C8, H4C9, PSMA6, SERPINA1*
	APEX1	Enzyme	2	Activated highly to M2	2	1.42E-06	9.00E-03	*APEX1, c-Src, HIF1A, SIRT1, TDG, temozolomide, TP53*	*ARPC1B, COL3A1, COL6A2, COX4I1, FTH1, GFAP, H2AC6, H2BC18, H4C15, HSPA5, HSPG2, LGALS3, LMAN2, MAPK1, NKX2-1, PCK2, RAP1B, RBM3, TINAGL1, TKT, UQCRC1*
	Cbp/p300	Group	2	Activated under both M1 and M2	2	1.74E-06	1.99E-02	*Cbp/p300, CTNNB1, EP300, ESR1, estrogen receptor, NCOA2, NCOA3, NFE2L2, NR3C2, PPARG, SMAD3, TP53*	*AKR1A1, ARPC1B, ATP6V0D1, CALD1, CLDN3, COL3A1, COL6A2, COX4I1, FTH1, GPD2, H4C15, H4C3, H4C6, HSPA5, HSPG2, LGALS3, LMAN2, MAPK1, MOGS, NKX2-1, PCK2, PSMA6, PSMD5, RBM3, SERPINA1, TINAGL1, UQCRC1*
	BUB1	Kinase	3	Activated highly to M2 but inhibited under NM	1.732	2.51E-07	3.80E-03	*APC, AURKB, BUB1, CTNNB1, GFAP, MTOR, PKM, TP53*	*AKR1A1, ARPC1B, COL3A1, COL6A2, COX4I1, FTH1, GFAP, H4C15, HSPA5, HSPG2, LGALS3, LMAN2, MAPK1, NKX2-1, PCK2, RBM3, SEC22B, SERPINA1, SND1, TINAGL1, TKT, UQCRC1*
	MAPK10	Kinase	2	Activated highly under M2 but inhibited under NM	1.732	1.40E-05	4.76E-02	*APP, CCND1, CDKN1A, JUN, L-serine, MAPK10, MAPKAPK3, NR3C1, TP53*	*ARPC1B, COL3A1, COL6A2, DPYSL2, FTH1, GFAP, H4C15, HSPA5, HSPG2, LGALS3, LMAN2, MAPK1, MOGS, PCK2, PRSS1, RBM3, SELENBP1, TINAGL1, TKT, TUFM, UQCRC1*
	HSF1	Transcription regulator	2	Inhibited commonly	−2.887	1.72E-11	1.00E-04	*CEBPB, EIF2A, ERK1/2, HSF1*	*COL3A1, COL6A2, GFAP, H4C1, H4C11, H4C12, H4C14, H4C15, H4C2, H4C4, H4C6, H4C8, H4C9, HSPA5, MAPK1, PCK2, PSMA6, SERPINA1*
	HNRNPK	Transcription regulator	2	Inhibited commonly	−2.887	1.06E-10	1.00E-04	*CEBPB, ERK1/2, HNRNPK, MAP2K1/2*	*COL3A1, GFAP, H4C1, H4C11, H4C12, H4C14, H4C15, H4C2, H4C4, H4C6, H4C8, H4C9, HSPA5, MAPK1, PSMA6, SERPINA1*

(Continued)

TABLE 2 | Continued

Module ID (module color)	Master regulator	Molecule type	Depth	Predicted activation state	Activation z-score	*p*-value of overlap	Network bias-corrected *p*-value	Participating regulators[a]	Target molecules in dataset[b]
	TGFB1	Growth factor	2	Upregurated under M2	1.342	1.31E-07	8.20E-03	*AKT1, EGFR, ERBB2, FOXO1, HIF1A, IKBKB, JUN, MTOR, MYB, MYC, NR2C2, RICTOR, SRF, STAT4, STK11, TGFB1, TP53*	*ARPC1B, ATP6V0D1, CALD1, CAPZA2, CDK5RAP3, CLDN3, COL3A1, COL6A2, DHRS2, DPYSL2, FBL, FTH1, GFAP, H2AC6, H4C15, H4C3, HBG2, HSPA5, HSPG2, LGALS3, LMAN2, LMNB1, TA4H, MAPK1, MOGS, NKX2-1, PGAM1, PSMA6, RBM3, SEC22B, SELENBP1, SERPINA1, SND1, SYPL1, TINAGL1, TUFM, UBE2L3, UQCRC1*

[a]*Participating regulators are regulators through which the upstream regulator molecule controls the expression of target molecules in the dataset.*
[b]*Target molecules in the dataset are molecules in our dataset whose expression is potentially controlled by an upstream regulator.*

The enriched pathways of the WM17 module involved neutrophil degranulation, immune response, and immune system (**Figure 6B**). The hub protein is the RAS-related protein Rab-5C (also known as L1880 or RAB5L). Rab-5C itself is one of the three isoforms of Rab-5, which is a master regulator of the endocytic pathway. The subnetworks related mostly to the immune system process are indicated in the pink dotted line 6 in **Figure 6A**. Protein FAM83D (also known as spindle protein CHICA), a probable proto-oncogene, plays a role in cell proliferation, growth, migration, and epithelial to mesenchymal transition (EMT) (56). Elevated FAM83D expressions were reported in several cancers including metastatic lung adenocarcinomas (57). Recently, Shi et al. suggested its oncogenic activity by regulating cell cycle in lung adenocarcinoma (58).

Comparative Analysis of Causal Networks Predicted by IPA

The ORA-based screening of the WGCNA modules was performed to capture clinically important modules and their upstream regulators, which reflect the disease mechanisms affected differentially under the different driver *EGFR* mutations in lung adenocarcinoma. Both upstream regulators and causal networks using IPA (http://www.ingenuity.com) software (43) were performed especially for the two selected modules, WM 10 and WM6, which were significantly associated with the Ex19del and L858R mutation (**Figures 5**, **6**). Causal networks predicted for these mutant protein modules included chemical drugs, transcriptional regulators, transmembrane receptors, growth factors, kinases, transporters, etc. **Table 2** summarizes top causal networks significant to each module ($|z\text{-value}| > 1.5$) representative under the *EGFR* L858R or Ex19del mutation status in the order of higher overlap significance, *p*-value. **Figure 7** presents the representative modules of master and participating regulators with the target mutant proteins differentially significant to the *EGFR* L858 or Ex19del mutation status.

Regarding the WM10 module associated with the M1 trait, the *EGFR* L858R mutation, *PARPBP*, *HOXA1*, and *APH-1* were highly activated or upregulated under the *EGFR* L858R mutation, whereas *AMOT* was highly inhibited under both L858R and Ex19del mutations. *PARPBP* encodes poly (ADP-ribose) polymerase-1 (PARP-1) binding protein, which plays a central role in DNA repair and the maintenance of genomic stability, regulating DNA repair, and negatively double-strand break repair via homologous recombination. Xu et al. reported that PARPBP expression was enhanced in lung adenocarcinoma tissues and correlated with poor prognosis in lung adenocarcinoma patients (59) and also that its high expression was closely correlated with pathologic stages, suggesting its utility as an independent predictor in lung adenocarcinoma patients. *HOXA1* encodes homeobox protein Hox-1F, a member of the Homeobox (HOX) transcription factor family. HOXA1 mRNA and protein expression levels were significantly upregulated in breast cancer, and its overexpression was associated with poor prognosis and tumor progression in breast cancer patients (60). Anterior

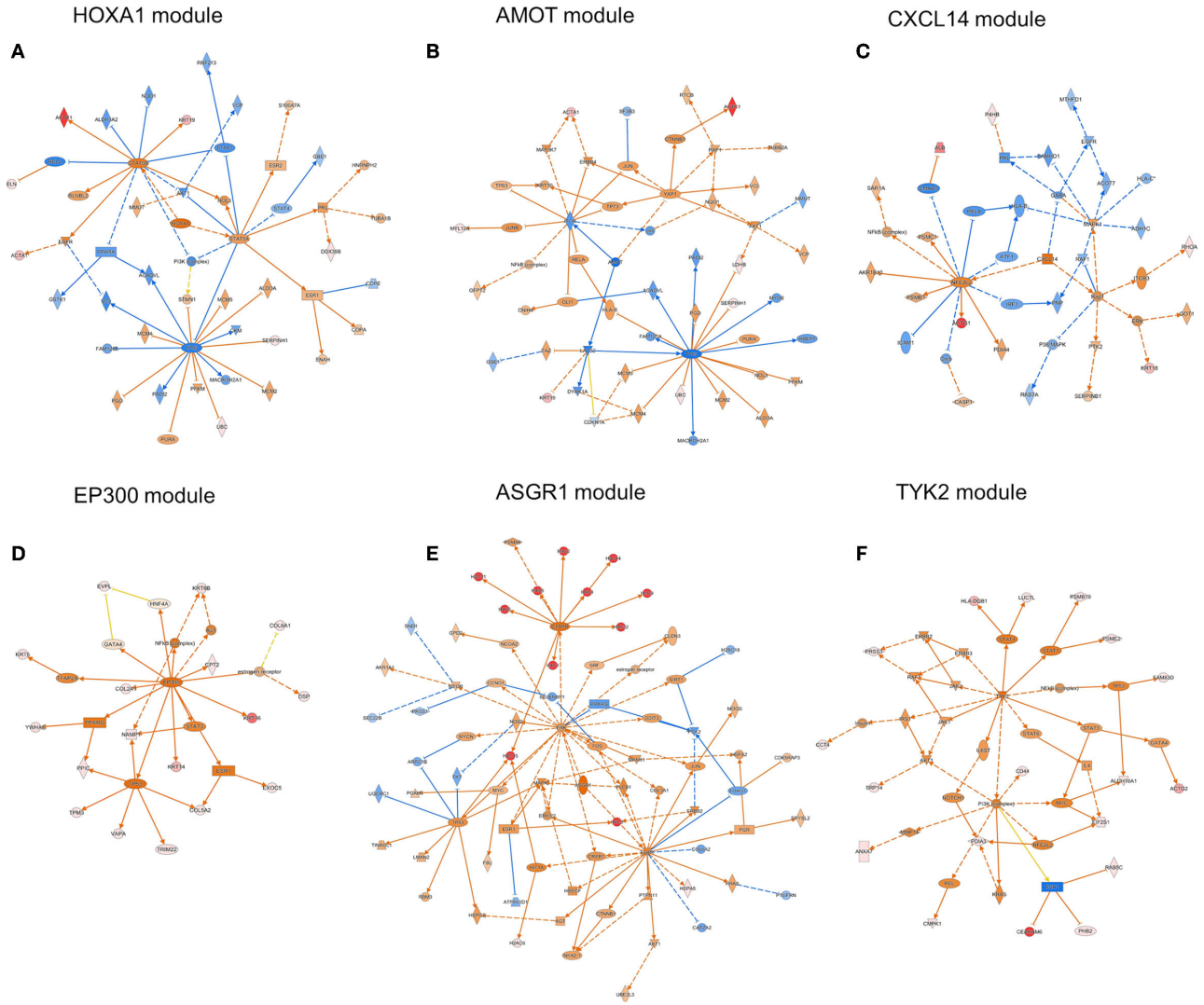


FIGURE 7 | The representative modules of master and participating regulators with the target mutant proteins differentially significant under the *EGFR* L858 or Ex19del mutation status, which were obtained by using IPA software. The modules of **(A)** *HOXA1*, **(C)** *CXCL14*, and **(D)** *EP300* were predicted to be highly activated and **(B)** *AMOT* highly inhibited in association with the L858R mutation status. The modules of **(E)** *ASGR1* and **(F)** *TYK2* were highly and/or differentially activated on the Ex19del mutation status. Node shapes indicate molecular types: triangle, kinase; square (dashed), growth factor; rectangle (horizontal), ligand-dependent nuclear receptor; rectangle (vertical), ion channel; diamond (vertical), enzyme; diamond (horizontal), peptidase; trapezoid, transporter; oval (horizontal), transcription regulator; oval (vertical), transmembrane receptor; double circle, complex; circle, other. Red or light red colors indicate highly or moderately increased expression of a mutant protein in the data set. Orange or light orange colors indicate the extent of confidence for predicted activation and the blue and light blue for predicted inhibition. Lines denote predicted relationships. A solid or dashed line indicates direct or indirect interaction, respectively. Orange indicates leading to activation; blue, leading to inhibition; yellow, findings inconsistent with the state of a downstream molecule; gray, an effect not predicted.

pharynx defective 1 (*APH1*) is the group *APH1A* and *APH1B,* which are the members of the gamma-secretase complex, comprising presenilin (*PSEN1* and *PSEN2*), anterior pharynx defective 1 (*APH1*), presenilin enhancer 2 (*PEN2*), and nicastrin. Gamma secretase substrates are known to include the four well-characterized mammalian Notch receptors (Notch1-4) and the five canonical transmembrane Notch ligands. Aberrant Notch activation drives development, tumorigenesis, and progression of lung cancer and is known to participate in resistance to anti-VEGF therapy (61). The inhibition of Notch activation by gamma-secretase inhibitors (GSIs) then could benefit NSCLC patients (62). Angiomotin (AMOT) and its related proteins, scaffold proteins, AMOT family proteins, were identified to have a strong interaction with the transcription factors Yes-associated protein (YAP) and TAZ (transcriptional coactivator with PDZ-binding motif) by tandem affinity purification (TAP) and mass spectrometry (63). Scaffold proteins angiomotin negatively regulated the transcription factors YAP and TAZ by preventing their nuclear translocation, suggesting a tumor-suppressing role of AMOT family proteins as components of the Hippo

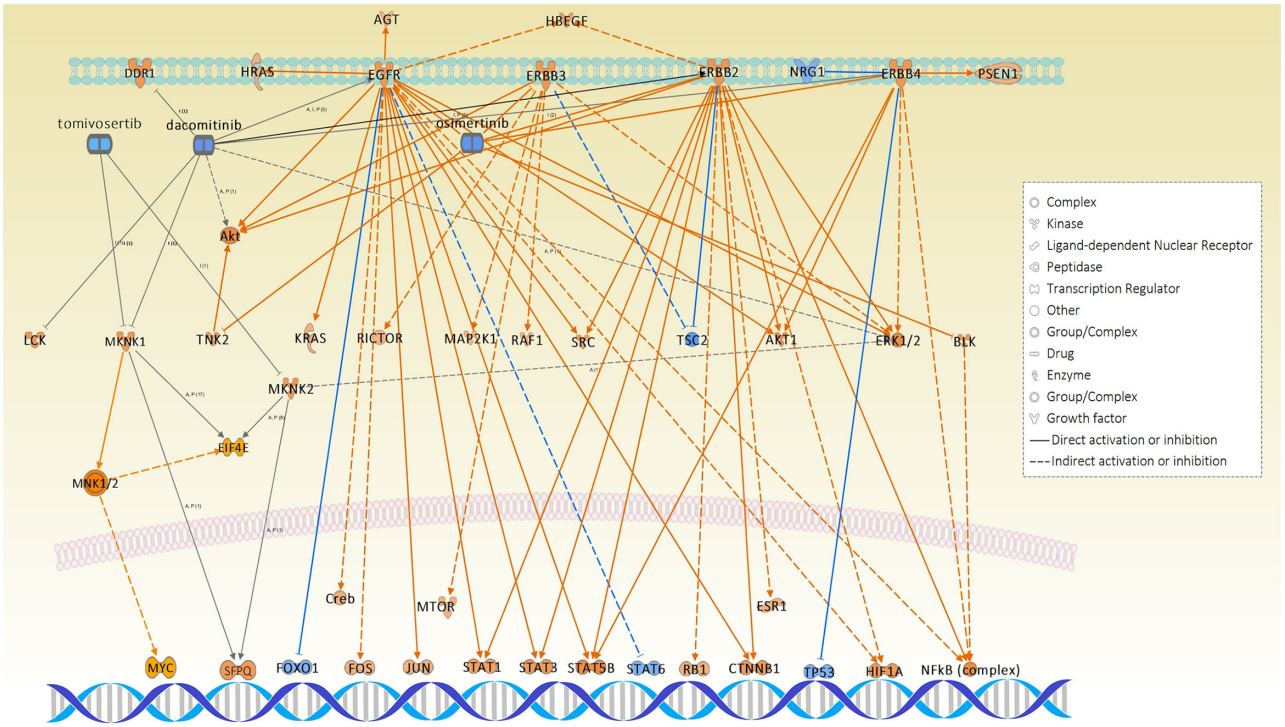


FIGURE 8 | The causal networks of downregulated osimertinib intervention and activated *MNK1/2*, which were predicted to be significant under L858R mutation status, illustrated together with the related inhibitors: dacomitinib and tomivosertib.

pathway. However, Hong reported the controversial results that AMOT may promote nuclear translocation of YAP and act as a transcriptional cofactor of the YAP-TEAD complex to facilitate the proliferation of epithelial cells and cancer development (64). It has been pointed out that the functional roles of AMOTs in different cancer types are controversial, highly depending on cell context (65).

Interestingly, among all causal networks predicted from the WM10 module, the downregulation of osimertinib intervention showed the highest significance in overlap *p*-value, in which *EGFR*, *ERBB2*, *ERBB3*, *ERBB4*, *BLK*, and *TNK* and their downstream pathways were maintained. In this study, we used FFPE tissue specimens collected from lung adenocarcinoma patients who did not receive any EGFR tyrosine kinase inhibitors, such as osimertinib. It has been reported that L858R-positive patients of NSCLC had a poor prognosis and difference in therapeutic outcome compared to Ex19del-positive patients (6). Moreover, the comparative IPA analysis predicted the *MNK1/2* causal network highly and differentially activated under the L858R mutation status (**Table S1**), which have been targeted by several chemical drug inhibitors for *EGFR* mutation-positive lung cancers. Those inhibitors, including dacomitinib, tomivosertib, BAY1143269, and ETC-1907206, have been developed for various types of *EGFR* mutation-positive cancers mainly including NSCLCs and various types of clinical trials are currently undergoing. Other top causal networks activated differentially under the *EGFR* L858R mutation included *max-myc* (complex), *MYC*, *F8*, *STK11*, and *RAD21*.

For the WM6 module associated with the M2 trait, the *EGFR* Ex19del mutation, *ASGR1* and *APEX1* were highly activated, and *BUB1*, *MAPK10*, and *TGFB1* were upregulated. *CEBPB* was activated commonly under all the traits, whereas *Cbp/p300* was activated under both Ex19del and L858R mutations. *ASGR1* encodes a subunit of the asialoglycoprotein receptor (*ASGR*) expressed in the extracellular region and a complex of the receptor and binding ligand is internalized. *ASGR* has been suggested to promote cancer metastasis by activating the EGFR–ERK pathway through interactions with counter-receptors on cancer cells, responding to endogenous lectins in the tumor microenvironment (66). *APEX1* (also known as *APE1*, *APX*, *HAP1*, and *REF1*) and encodes DNA-apurinic/apyrimidinic (AP) site endonuclease (protein names, such as APEN, APE-1, and REF-1), which plays a central role in the cellular response to oxidative stress, in which its two major activities are DNA repair and redox regulation of transcriptional factors. The elevated levels of APEX1 have been reported in several cancers, including lung cancer (67), and also to be associated with resistance to chemotherapy and radiotherapy in some cancers (68). *MAPK10* encodes mitogen-activated protein kinase 10 (also known as stress-activated protein kinase JNK3), which is involved in a wide variety of cellular processes, including stress response, proliferation, differentiation, transcription regulation, and development. *MAPK10* functions as a tumor suppressor and the deletion of this proapoptotic gene would favor the survival and proliferation of cancer cells (69). *BUB1* encodes mitotic checkpoint serine/threonine-protein kinase BUB1 or budding

uninhibited by benzimidazoles 1 (Bub1), which is required for chromosome alignment and resolution of spindle attachment errors but does not play a major role in the spindle assembly checkpoint (SAC) activity. Overexpression of Bub1 in breast cancer is associated with a poor clinical prognosis (70). Recent tumor xenograft studies suggested that the Bub1 kinase inhibitor BAY 1816032 in combination with taxanes or PARP inhibitors enhanced their efficacy and suppressed the development of therapy resistance (71). *CEBPB* encodes CCAAT/enhancer-binding protein beta (C/EBP beta), which is important in the regulation of genes involved in immune and inflammatory responses. C/EBP beta induces elevated IL-6 expression levels frequently observed in human lung adenocarcinomas (72) and interacts with peroxisome proliferator-activated receptor-gamma (*PPARG*) involved in pathways of transcriptional misregulation in cancer (73). The study using the inducible EGFR T790M-L858R transgenic mouse models suggested that C/EBP beta is dispensable for lung tumorigenesis in *EGFR*-driven murine lung cancer (73).

DISCUSSION

Outcomes of lung adenocarcinoma patients receiving EGFR TKIs were reported to be affected depending on the types of *EGFR* gatekeeper mutation (6, 74), which are serious clinical challenges. Targeting disease-associated dual core networks rather than targeting a single protein (gene) as in conventional approaches is expected to greatly improve the outcomes of individual patients, such as efficacy and safety, in line with the concept of precision medicine. Such a concept, so-called network pharmacology, was first proposed by Hopkins (75), which aims to induce synthetic lethality by targeting dual hub molecules involved in different disease core networks. We have first conducted a mutant proteomic analysis for clinical tissue specimens of 36 lung adenocarcinoma patients who harbored distinct *EGFR* mutations, Ex21 L858R, Ex19del, and no L858R/Ex19del. Disease-related network modules are elucidated from mutant protein expression data sets, which would be potentially associated with the activation of downstream and/or upstream networks affected under distinct *EGFR* mutations. In particular, this study focuses on influence in disease-related networks of lung adenocarcinoma, which would take place under the L858R mutation. Our analytical workflow combining WGCNA with ORA-screening identified several mutant protein modules significantly overlapping with upregulated mutant proteins under the *EGFR* L858R mutation.

Our goal with the present study was to apply an unbiased bioinformatic method to characterize the mutant profiles of detectable SAAVs after filtering with stringent criteria of database identifications in pathologically well-described patient samples. Mass spectrometry-based proteomic data is widely recognized as an information-rich source of uniquely expressed proteoforms, but tandem mass spectra interpretation is dependent on fragmentation efficiency and identification strategies. Because the number of subjects was limited in each patient group, we presented quality control data in the **Supplementary Material** demonstrating the overall homogeneity of the mass spectrometric data due to low technical variability of sample preparation and data acquisition. Careful interpretation of the findings highlighted potential differences between phenotypes, which suggests that different oncogenic driver *EGFR* mutations would affect activation or inactivation of their downstream disease-related molecular networks, which are often associated with protein mutations.

Surprisingly, the OPLS-DA performed for identified mutant proteins demonstrated profound differences in distance among the different *EGFR* mutation groups, L858R, Ex19del, and no L858R/Ex19del, suggesting that cancer cells harboring L858R or Ex19del emerge from cellular origins differently from L858R/Ex19del-negative cells (**Figure 2B**). Aberrant cells would, thus, emerge as a subpopulation of tumor cells of genetic intratumor heterogeneity, which would rapidly grow and predominantly survive by disrupting the tumor environment. To confirm our observation, a further large-scale investigation with genomic alteration analysis by next-generation sequencing (NGS) is required.

The pathways of the carboxylic acid metabolic process, cell cycle, developmental biology, and immune system were centrally associated under the L858R mutation. The top IPA causal networks predicted for the representative mutant protein module-WM10 were associated with the regulation of DNA repair, cancer development, tumorigenesis, and maintenance of genomic stability as well as therapeutic resistance. Interestingly, the downregulation of osimertinib intervention showed the highest significance rank in overlap among all causal networks predicted from the WM10 module (**Table 2**). This finding might suggest the potential usefulness of osimertinib to be revisited for the L858R-positive patients of lung adenocarcinoma. Both the causal networks of osimertinib intervention and *MNK1/2* identified significantly and differentially, respectively, may evidence disease mechanisms associated with *EGFR* mutation-positive lung adenocarcinoma (**Figure 8**).

The pathways of cellular responses to stress, mitotic prophase, cell proliferation, growth, migration, epithelial to mesenchymal transition (EMT), and immune system process were mostly involved under the Ex19del mutation. The IPA causal networks elucidated for the representative mutant protein module, WM6, seem to be associated dominantly with the EGFR–ERK pathway. The pathways related to the Hippo pathway and tumorigenesis were commonly involved under both L858R and Ex19del mutations.

The limitations of this study are as follows: first, the number of patients examined is limited to be 36, which was attributed to collect the homogeneous tumor-derived samples with the best effort. Second, genomic alteration analysis was not conducted for the same samples.

In conclusion, we successfully applied WGCNA combined with ORA-based protein screening to clinical mutant proteomic data sets from 36 patients of lung adenocarcinoma. The proteomic discovery method detecting mutant proteoforms has revealed specific profiles distinguishing the phenotypically characterized patient groups. Our results could confirm the usefulness of mutant proteomics to identify activated or

inactivated disease-related mutant protein networks affected under distinct *EGFR* mutations. Verification and quantitative analysis of these molecular features in an independent patient cohort are yet to be undertaken by either using targeted proteomics or RNAseq and combining the resulting data in a systems biology approach. Additionally, our findings may help in the development of therapeutic strategies to improve patient outcomes. Differences in mutant proteomes between L858R and Ex19del mutation cells help to demonstrate the difference in efficacy of various EGFR-TKIs. Further verifications with a greater number of patient samples and targeted analysis of mutant proteoforms throughout the cohorts are planned in follow-up studies to achieve a better understanding of the expression profiles of SAAVs in phenotypic groups and establish a relationship between the detected networks in connection to disease progression.

DATA AVAILABILITY STATEMENT

The unfiltered mass spectrometry data sets generated and analyzed in this study have been deposited in the ProteomeXchange (http://proteomecentral.proteomexchange.org) and jPOST with the data set identifiers PXD015862 and JPST000687, respectively.

ETHICS STATEMENT

FFPE tumor tissue blocks from 36 surgical specimens of lung adenocarcinomas with known *EGFR* mutation statuses were obtained without patient identifiers from St. Marianna University School of Medicine Hospital. Informed consent was obtained from all participating subjects, and the protocol was approved by the institutional review board of St. Marianna University School of Medicine (approval no. 3569) and was conducted in accordance with the Helsinki Declaration.

AUTHOR CONTRIBUTIONS

TN and ÁV conceptualized this study, designed the bioinformatics methodology, wrote the main manuscript text, prepared **Figures 1**–**7** and **Supplementary Information Files**, and wrote the first draft of the manuscript. ÁV performed the identification of mutant protein from the clinical proteome raw data sets. TN analyzed the mutant proteome data sets by using weighted gene coexpression network analysis (WGCNA), over representative analysis, and the ingenuity pathway analysis software. TN, HN, HK, and HS initiated and managed the collaboration. All authors reviewed the manuscript.

FUNDING

The authors declare that this study received funding from AstraZeneca K.K. on this research project under the Externally Sponsored Research ESR-16-12243. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01494/full#supplementary-material

Supplementary Information File 1 | Search summaries, including score distributions and statistical data are presented using the PEAKS reporting function.

Supplementary Information File 2 | The complete lists of identified mutant protein sequences after filtration.

Figure S1 | Geneontology (GO) analysis of the identified proteins to the three traits-M1 (L858R), M2 (Ex19del), and NM (no L858R/Ex19del). **(A)** GO Biological process. 1, developmental process (GO:0032502); 2, multicellular organismal process (GO:0032501); 3, cellular process (GO:0009987); 4, reproduction (GO:0000003); 5, cell population proliferation (GO:0008283); 6, localization (GO:0051179); 7, reproductive process (GO:0022414); 8, multiorganism process (GO:0051704); 9, biological adhesion (GO:0022610); 10, immune system process (GO:0002376); 11, cellular component organization or biogenesis (GO:0071840); 12, biological regulation (GO:0065007); 13, growth (GO:0040007); 14, signaling (GO:0023052); 15, metabolic process (GO:0008152); 16, response to stimulus (GO:0050896); 17, pigmentation (GO:0043473); 18, behavior (GO:0007610); 19, locomotion (GO:0040011). **(B)** GO Molecular function. 1, translation regulator activity (GO:0045182); 2, transcription regulator activity (GO:0140110); 3, molecular transducer activity (GO:0060089); 4, binding (GO:0005488); 5, structural molecule activity (GO:0005198); 6, molecular function regulator (GO:0098772); 7, catalytic activity (GO:0003824); 8, transporter activity (GO:0005215). **(C)** GO Cellular component. 1, synapse part (GO:0044456); 2, membrane part (GO:0044425); 3, membrane (GO:0016020); 4, synapse (GO:0045202); 5, organelle part (GO:0044422); 6, extracellular region part (GO:0044421); 7, cell junction (GO:0030054); 8, membrane-enclosed lumen (GO:0031974); 9, protein-containing complex (GO:0032991); 10, supramolecular complex (GO:0099080); 11, extracellular region (GO:0005576); 12, cell (GO:0005623); 13, cell part (GO:0044464); 14, organelle (GO:0043226). **(D)** GO Protein class. 1, extracellular matrix protein (PC00102); 2, cytoskeletal protein (PC00085); 3, transporter (PC00227); 4, scaffold/adaptor protein (PC00226); 5, cell adhesion molecule (PC00069); 6, nucleic acid binding protein (PC00171); 7, intercellular signal molecule (PC00207); 8, protein-binding activity modulator (PC00095); 9, calcium-binding protein (PC00060); 10, gene-specific transcriptional regulator (PC00264); 11, defense/immunity protein (PC00090); 12, translational protein (PC00263); 13, metabolite interconversion enzyme (PC00262); 14, protein modifying enzyme (PC00260); 15, chromatin/chromatin-binding, or -regulatory protein (PC00077); 16, transfer/carrier protein (PC00219); 17, membrane traffic protein (PC00150); 18, chaperone (PC00072); 19, cell junction protein (PC00070); 20, structural protein (PC00211); 21, storage protein (PC00210); 22, transmembrane signal receptor (PC00197).

Figure S2 | Relationship between module eigen proteins and the L858R and Ex19del mutations in the *EGFR* gene. Each row in the embedded table represents weighted gene coexpression network analysis results for each module. The first and second columns in the table represent module ID and color name of the module. The third column represents the number of proteins in each module. The fourth, fifth, and sixth (seventh, eighth, and ninth) columns indicate the correlation coefficients (*p*-values of the correlation coefficients) between the corresponding modules and the clinical traits. The table is color-coded by correlation coefficient according to the color legend on the right side of the figure. The intensity and direction of the correlations are indicated on the right side of the heat map (red, positive correlation; blue, negative correlation). *p*-values (<0.10) are highlighted in red.

Figure S3 | R_{SC} values between M1 and M2 calculated for proteins identified (X-axis). Mutant proteins upregulated with 2-fold changes for M1 ($R_{SC} \geq 1$) and M2 ($R_{SC} \leq -1$) are denoted.

Table S1 | The comparative analysis results of causal networks predicted for mutant proteins expressed commonly (see Venn map in **Figure 2A**). *MNK1/2*, *Max-Myc*, MYC, *XBP1*, *BTG2*, *F8*, *STK11*, and *RAD21* were highly activated (*z*-score > 2.5) and differentially under M1 (L858R).

REFERENCES

1. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med.* (2004) 350:2129–39. doi: 10.1056/NEJMoa040938
2. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science.* (2004) 304:1497–500. doi: 10.1126/science.1099314
3. Kobayashi Y, Mitsudomi T. Not all epidermal growth factor receptor mutations in lung cancer are created equal: perspectives for individualized treatment strategy. *Cancer Sci.* (2016) 107:1179–86. doi: 10.1111/cas.12996
4. Kobayashi S, Boggon TJ, Dayaram T, Jänne PA, Kocher O, Meyerson M, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med.* (2005) 352:786–92. doi: 10.1056/NEJMoa044238
5. Mok TS, Wu Y-L, Ahn M-J, Garassino MC, Kim HR, Ramalingam SS, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med.* (2017) 376:629–40. doi: 10.1056/NEJMoa1612674
6. Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N Engl J Med.* (2018) 378:113–25. doi: 10.1056/NEJMoa1713137
7. Ramalingam SS, Vansteenkiste J, Planchard D, Cho BC, Gray JE, Ohe Y, et al. Osimertinib vs comparator EGFR-TKI as first-line treatment for EGFRm advanced NSCLC (FLAURA): final overall survival analysis. *Ann Oncol.* (2019) 30:v914–15. doi: 10.1093/annonc/mdz394.076
8. Forcella M, Oldani M, EpistolioS, Freguia S, Monti E, Fusi P, et al. Non-small cell lung cancer (NSCLC), EGFR downstream pathway activation and TKI targeted therapies sensitivity: effect of the plasma membrane-associated NEU3. *PLoS ONE.* (2017) 12:e0187289. doi: 10.1371/journal.pone.0187289
9. Torres AF, Nogueira C, Magalhaes J, Costa IS, Aragao A, Neto AG, et al. Expression of EGFR and molecules downstream to PI3K/Akt, Raf-1-MEK-1-MAP (Erk1/2), and JAK (STAT3) pathways in invasive lung adenocarcinomas resected at a single institution. *Anal Cell Pathol (Amst).* (2014) 2014:352925. doi: 10.1155/2014/352925
10. Azevedo AP, Silva S, Rueff J. Non-receptor Tyrosine Kinases Role and Significance in Hematological Malignancies. (2019). doi: 10.5772/intechopen.84873
11. Han G, Feng J, Peng M, Verma V, Bi J, Song Q. EGFR overexpression and mutations lead to a change in biological characteristics of human lung adenocarcinoma cells. *Int J Radiation Oncol.* (2017) 99:E594. doi: 10.1016/j.ijrobp.2017.06.2031
12. Wee P, Wang Z. Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers.* (2017) 9:52. doi: 10.3390/cancers9050052
13. Morgillo F, Della Corte CM, Fasano M, Ciardiello F. Mechanisms of resistance to EGFR-targeted drugs: lung cancer. *ESMO Open.* (2016) 1:e000060. doi: 10.1136/esmoopen-2016-000060
14. Liu Q, Yu S, Zhao W, Qin S, Chu Q, Wu K. EGFR-TKIs resistance via EGFR-independent signaling pathways. *Mol Cancer.* (2018) 17:53. doi: 10.1186/s12943-018-0793-1
15. Peng S, Wang R, Zhang X, Ma Y, Zhong L, Li K, et al. EGFR-TKI resistance promotes immune escape in lung cancer via increased PD-L1 expression. *Mol Cancer.* (2019) 18:165. doi: 10.1186/s12943-019-1073-4
16. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer.* (2009) 9:798–809. doi: 10.1038/nrc2734
17. Yan S, Li Z, Thiele CJ. Inhibition of STAT3 with orally active JAK inhibitor, AZD1480, decreases tumor growth in neuroblastoma and pediatric sarcomas *in vitro* and *in vivo*. *Oncotarget.* (2013) 4:433–45. doi: 10.18632/oncotarget.930
18. Wu K, Chang Q, Lu Y, Qiu P, Chen B, Thakur C, et al. Gefitinib resistance resulted from STAT3-mediated Akt activation in lung cancer cells. *Oncotarget.* (2013) 4:2430–8. doi: 10.18632/oncotarget.1431
19. Jin Y, Bao H, Le X, Fan X, Tang M, Shi X, et al. Distinct co-acquired alterations and genomic evolution during TKI treatment in non-small-cell lung cancer patients with or without acquired T790M mutation. *Oncogene.* (2020) 39:1846–59. doi: 10.1038/s41388-019-1104-z
20. Harrison PT, Vyse S, Huang PH. Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer. *Semin Cancer Biol.* (2020) 61:167–79. doi: 10.1016/j.semcancer.2019.09.015
21. Yates JR. The revolution and evolution of shotgun proteomics for large-scale proteome analysis. *J Am Chem Soc.* (2013) 135:1629–40. doi: 10.1021/ja3094313
22. Tong J, Taylor P, Moran MF. Proteomic analysis of the epidermal growth factor receptor (EGFR) interactome and post-translational modifications associated with receptor endocytosis in response to EGF and stress. *Mol Cell Proteomics.* (2014) 13:1644–58. doi: 10.1074/mcp.M114.038596
23. Putri DU, Chiumia FK, Jheng YT, Han CL. The role of proteomics for dissecting aberrant molecular signaling pathways upon Egfr-Tki treatments in non-small cell lung cancer. *Proteomics Bioinform Curr Res.* (2019) 1:4–16.
24. Zhang X, Belkina N, Jacob HK, Maity T, Biswas R, Venugopalan A, et al. Identifying novel targets of oncogenic EGF receptor signaling in lung cancer through global phosphoproteomics. *Proteomics.* (2015) 15:340–55. doi: 10.1002/pmic.201400315
25. Zhang X, Maity T, Kashyap MK, Bansal M, Venugopalan A, Singh S, et al. Quantitative tyrosine phosphoproteomics of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor-treated lung adenocarcinoma cells reveals potential novel biomarkers of therapeutic response. *Mol Cell Proteomics.* (2017) 16:891–910. doi: 10.1074/mcp.M117.067439
26. Végvári Á. Mutant proteogenomics. In: Végvári Á, editor. *Proteogenomics*. Cham: Springer (2016). p. 77–91. doi: 10.1007/978-3-319-42316-6_6
27. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* (2007) 449:851–61. doi: 10.1038/nature06258
28. Nishimura T, Nakamura H, Végvári Á, Marko-Varga G, Furuya N, Saji H. Current status of clinical proteogenomics in lung cancer. *Expert Rev Proteomics.* (2019) 16:761–72. doi: 10.1080/14789450.2019.1654861
29. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science.* (2007) 318:1108–13. doi: 10.1126/science.1145720
30. Sun T, Zhou Y, Yang M, Hu Z, Tan W, Han X, et al. Functional genetic variations in cytotoxic T-lymphocyte antigen 4 and susceptibility to multiple types of cancer. *Cancer Res.* (2008) 68:7025–34. doi: 10.1158/0008-5472.CAN-08-0806
31. Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic variation in human gene expression. *Science.* (2002) 297:1143. doi: 10.1126/science.1072545
32. Végvári A, Sjödin K, Rezeli M, Malm J, Lilja H, Laurell T, et al. Identification of a novel proteoform of prostate specific antigen (SNP-L132I) in clinical samples by selective reaction monitoring. *Mol Cell Proteomics.* (2013) 12:2761–73. doi: 10.1074/mcp.M113.028365
33. Wang Q, Chaerkady R, Wu J, Hwang HJ, Papadopoulos N, Kopelovich L, et al. Mutant proteins as cancer-specific biomarkers. *Proc Natl Acad Sci USA.* (2011) 108:2444–9. doi: 10.1073/pnas.1019203108
34. Kawamura T, Nomura M, Tojo H, Fujii K, Hamasaki H, Mikami S, et al. Proteomic analysis of laser microdissected paraffin-embedded tissues: (1) Stage-related protein candidates upon non-metastatic lung adenocarcinoma. *J Proteomics.* (2010) 73:1089–99. doi: 10.1016/j.jprot.2009.11.011
35. Fujii K, Miyata Y, Takahashi I, Koizumi H, Saji H, Hoshikawa M, et al. Differential proteomic analysis between small cell lung carcinoma (SCLC) and pulmonary carcinoid tumors reveals molecular signatures for malignancy in lung cancer. *Proteomics Clin Appl.* (2018) 12:e1800015. doi: 10.1002/prca.201800015
36. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* (2008) 9:559. doi: 10.1186/1471-2105-9-559

37. Tang Y, Ke ZP, Peng YG, Cai PT. Coexpression analysis reveals key gene modules and pathways of human coronary heart disease. *J Cell Biochem*. (2018) 119:2102–9. doi: 10.1002/jcb.26372
38. Nakamura H, Fujii K, Gupta V, Hata H, Koizumu H, Hoshikawa M, et al. Identification of key modules and hub genes for small-cell lung carcinoma and large-cell neuroendocrine lung carcinoma by weighted gene co-expression network analysis of clinical tissue-proteomes. *PLoS ONE*. (2019) 14:e0217105. doi: 10.1371/journal.pone.0217105
39. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol*. (2015) 10:1243–60. doi: 10.1097/JTO.0000000000000630
40. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, et al. Comparison of label-free methods for quantifying human proteins by discovery proteomics. *Mol Cell Proteomics*. (2005) 4:1487–502. doi: 10.1074/mcp.M500084-MCP200
41. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, et al. PEAKS DB: *de novo* sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics*. (2012) 11:M111.010587. doi: 10.1074/mcp.M111.010587
42. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J. et al. STRING v11:protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. (2019) 47:D607–13. doi: 10.1093/nar/gky1131
43. Krämer A, Green J, Pollard J Jr., Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. (2014) 30:523–30. doi: 10.1093/bioinformatics/btt703
44. Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc*. (2019) 14:703–21. doi: 10.1038/s41596-019-0128-8
45. Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemometrics*. (2006) 20:341–51. doi: 10.1002/cem.1006
46. Khomtchouk BB, Hennessy JR, Wahlestedt C. shinyheatmap: Ultra fast low memory heatmap web interface for big data genomics. *PLoS ONE*. (2017) 12:e0176334. doi: 10.1371/journal.pone.0176334
47. Yan X, Chang J, Sun R, Meng X, Wang W, Zeng L, et al. DHX9 inhibits epithelial-mesenchymal transition in human lung adenocarcinoma cells by regulating STAT3. *Am J Transl Res*. (2019) 11:4881–94.
48. Thompson E. SAMHD1 is a novel biomarker and therapeutic target for radiation therapy and PARP inhibition in breast cancer. *Cancer Res*. (2019) 79:Abstract nr 3497. doi: 10.1158/1538-7445.AM2019-3497
49. Rentoft M, Lindell K, Tran P, Chabes AL, Buckland RJ, Watt DL, et al. Heterozygous colon cancer-associated mutations of SAMHD1 have functional significance. *Proc Natl Acad Sci USA*. (2016) 113:4723–28. doi: 10.1073/pnas.1519128113
50. Herold N, Rudd SG, Sanjiv K, Kutzner J, Myrberg IH, Paulin CBJ, et al. With me or against me: tumor suppressor and drug resistance activities of SAMHD1. *Exp Hematol*. (2017) 52:32–9. doi: 10.1016/j.exphem.2017.05.001
51. Zhuo DX, Zhang XW, Jin B, Zhang Z, Xie BS, Wu CL, et al. CSTP1, a novel protein phosphatase, blocks cell cycle, promotes cell apoptosis, and suppresses tumor growth of bladder cancer by directly dephosphorylating Akt at Ser473 site. *PLoS ONE*. (2013) 8:e65679. doi: 10.1371/journal.pone.0065679
52. Wu J1, Jiang H, Luo S, Zhang M, Zhang Y, Sun F, et al. Caspase-mediated cleavage of C53/LZAP protein causes abnormal microtubule bundling and rupture of the nuclear envelope. *Cell Res*. (2013) 23:691–704. doi: 10.1038/cr.2013.36
53. Tan D, Li Q, Deeb G, Ramnath N, Slocum HK, Brooks J, et al. Thyroid transcription factor-1 expression prevalence and its clinical implications in non-small cell lung cancer: a high-throughput tissue microarray and immunohistochemistry study. *Hum Pathol*. (2003) 34:597–604. doi: 10.1016/S0046-8177(03)00180-1
54. Tanaka H, Yanagisawa K, Shinjo K, Taguchi A, Maeno K, Tomida S, et al. Lineage-specific dependency of lung adenocarcinomas on the lung development regulator TTF-1. *Cancer Res*. (2007) 67:6007–11. doi: 10.1158/0008-5472.CAN-06-4774
55. Zu YF, Wang XC, Chen Y, Wang JY, Liu X, Li X, et al. Thyroid transcription factor 1 represses the expression of Ki-67 and induces apoptosis in non-small cell lung cancer. *Oncol Rep*. (2012) 28:1544–50. doi: 10.3892/or.2012.2009
56. Wang Z, Liu Y, Zhang P, Zhang W, Wang W, Curr K, et al. FAM83D promotes cell proliferation and motility by downregulating tumor suppressor gene FBXW7. *Oncotarget*. (2013) 4:2476–86. doi: 10.18632/oncotarget.1581
57. Inamura K, Shimoji T, Ninomiya H, Hiramatsu M, Okui M, Satoh Y, et al. A metastatic signature in entire lung adenocarcinomas irrespective of morphological heterogeneity. *Hum Pathol*. (2007) 38:702–9. doi: 10.1016/j.humpath.2006.11.019
58. Shi R, Sun J, Sun Q, Zhang Q, Xia W, Dong G, et al. Upregulation of FAM83D promotes malignant phenotypes of lung adenocarcinoma by regulating cell cycle. *Am J Cancer Res*. (2016) 6:2587–98.
59. Xu D, Tao Z, Tang X, He JK. Poly (ADP-ribose) polymerase-1 binding protein facilitates lung adenocarcinoma cell proliferation and correlates with poor prognosis. *Ann Clin Lab Sci*. (2019) 49:574–80.
60. Liu J, Liu J, Lu X. HOXA1 upregulation is associated with poor prognosis and tumor progression in breast cancer. *Exp Ther Med*. (2019) 17:1896–902. doi: 10.3892/etm.2018.7145
61. Li JL, Sainson RC, Oon CE, Turley H, Leek R, Sheldon H, et al. DLL4-Notch signaling mediates tumor resistance to anti-VEGF therapy *in vivo*. *Cancer Res*. (2011) 71:6073–83. doi: 10.1158/0008-5472.CAN-11-1704
62. Pine SR. Rethinking gamma-secretase inhibitors for treatment of non-small-cell lung cancer: is notch the target? *Clin Cancer Res*. (2018) 24:6136–41. doi: 10.1158/1078-0432.CCR-18-1635
63. Zhao B1, Li L, Lu Q, Wang LH, Liu CY, Lei Q, et al. Angiomotin is a novel Hippo pathway component that inhibits YAP oncoprotein. *Genes Dev*. (2011) 25:51–63. doi: 10.1101/gad.2000111
64. Hong W. Angiomotin'g YAP into the nucleus for cell proliferation and cancer development. *Sci. Signal*. (2013) 6:pe27. doi: 10.1126/scisignal.2004573
65. Huang T, Zhou Y, Zhang J, Cheng ASL, Yu J, To KF, et al. The physiological role of Motin family and its dysregulation in tumorigenesis. *J Transl Med*. (2018) 16:98. doi: 10.1186/s12967-018-1466-y
66. Ueno S, Mojic M, Ohashi Y, Higashi N, Hayakawa Y, Irimura T. Asialoglycoprotein receptor promotes cancer metastasis by activating the EGFR-ERK pathway. *Cancer Res*. (2011) 71:6419–27. doi: 10.1158/0008-5472.CAN-11-1773
67. Kakolyris S, Giatromanolaki A, Koukourakis M, Kaklamanis L, Kanavaros P, Hickson ID, et al. Nuclear localization of human AP endonuclease 1 (HAP1/Ref-1) associates with prognosis in early operable non-small cell lung cancer (NSCLC). *J Pathol*. (1999) 189:351–7. doi: 10.1002/(SICI)1096-9896(199911)189:3<351::AID-PATH435>3.0.CO;2-1
68. Tell G, Quadrifoglio F, Tiribelli C, Kelley MR. The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid Redox Signal*. (2009) 11:601–20. doi: 10.1089/ars.2008.2194
69. Xie Y, Liu Y, Fan X, Zhang L, Li Q, Li S, et al. MicroRNA-21 promotes progression of breast cancer via inhibition of mitogen-activated protein kinase10 (MAPK10). *Biosci Rep*. (2019). doi: 10.1042/BSR20181000. [Epub ahead of print].
70. Takagi K, Miki Y, Shibahara Y, Nakamura Y, Ebata A, Watanabe M, et al. BUB1 immunolocalization in breast carcinoma: its nuclear localization as a potent prognostic factor of the patients. *Horm Cancer*. (2013) 4:92–102. doi: 10.1007/s12672-012-0130-x
71. Siemeister G, Mengel A, Fernández-Montalván AE, Bone W, Schröder J, Zitzmann-Kolbe S, et al. Inhibition of BUB1 kinase by BAY 1816032 sensitizes tumor cells toward taxanes, ATR, and PARP inhibitors *in vitro* and *in vivo*. *Clin Cancer Res*. (2019) 25:1404–14. doi: 10.1158/1078-0432.CCR-18-0628
72. Meng X, Lu P, Bai H, Xiao P, Fan Q. Transcriptional regulatory networks in human lung adenocarcinoma. *Mol Med Rep*. (2012) 6:961–6. doi: 10.3892/mmr.2012.1034

73. Cai Y, Hirata A, Nakayama S, VanderLaan PA, Levantini E, Yamamoto M, et al. CCAAT/enhancer binding protein β is dispensable for development of lung adenocarcinoma. *PLoS ONE*. (2015) 10:e0120647. doi: 10.1371/journal.pone.0120647
74. Saito H, Fukuhara T, Furuya N, Watanabe K, Sugawara S, Iwasawa S, et al. Erlotinib plus bevacizumab versus erlotinib alone in patients with EGFR-positive advanced non-squamous non-small-cell lung cancer (NEJ026): interim analysis of an open-label, randomized, multicentre, phase 3 trial. *Lancet Oncol*. (2019) 20:625–35. doi: 10.1016/S1470-2045(19)30035-X
75. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. (2008) 4:682–90. doi: 10.1038/nchembio.118

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 © 2020 Nishimura, Végvári, Nakamura, Kato and Saji. 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 Oncology

ORIGINAL RESEARCH
published: 26 August 2020
doi: 10.3389/fonc.2020.01600

Comprehending Meningioma Signaling Cascades Using Multipronged Proteomics Approaches & Targeted Validation of Potential Markers

*Shuvolina Mukherjee[1], Deeptarup Biswas[1], Rucha Gadre[1], Pooja Jain[2], Nelofer Syed[3], Julianna Stylianou[3], Qingyu Zeng[3], Anita Mahadevan[4], Sridhar Epari[5], Prakash Shetty[6], Aliasgar Moiyadi[6], Graham Roy Ball[7] and Sanjeeva Srivastava[1]**

[1] Proteomics Lab, Department of Biosciences and Bioengineering, IIT Bombay, Mumbai, India, [2] Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London, United Kingdom, [3] Division of Brain Sciences, Department of Medicine, Imperial College London, London, United Kingdom, [4] Department of Neuropathology, Human Brain Tissue Repository (Brain Bank), NIMHANS, Bengaluru, India, [5] Department of Pathology, Tata Memorial Centre, Mumbai, India, [6] Department of Neurosurgery, Tata Memorial Centre, Mumbai, India, [7] School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Germain Gillet,
Université Claude Bernard Lyon 1, France
Parvez Khan,
University of Nebraska Medical Center, United States

***Correspondence:**
Sanjeeva Srivastava
sanjeeva@iitb.ac.in

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 31 March 2020
Accepted: 23 July 2020
Published: 26 August 2020

Citation:
Mukherjee S, Biswas D, Gadre R, Jain P, Syed N, Stylianou J, Zeng Q, Mahadevan A, Epari S, Shetty P, Moiyadi A, Roy Ball G and Srivastava S (2020) Comprehending Meningioma Signaling Cascades Using Multipronged Proteomics Approaches & Targeted Validation of Potential Markers. Front. Oncol. 10:1600. doi: 10.3389/fonc.2020.01600

Meningiomas are one of the most prevalent primary brain tumors. Our study aims to obtain mechanistic insights of meningioma pathobiology using mass spectrometry-based label-free quantitative proteome analysis to identifying druggable targets and perturbed pathways for therapeutic intervention. Label-free based proteomics study was done from peptide samples of 21 patients and 8 non-tumor controls which were followed up with Phosphoproteomics to identify the kinases and phosphorylated components of the perturbed pathways. *In silico* approaches revealed perturbations in extracellular matrix remodeling and associated cascades. To assess the extent of influence of Integrin and PI3K-Akt pathways, we used an Integrin Linked Kinase inhibitor on patient-derived meningioma cell line and performed a transcriptomic analysis of the components. Furthermore, we designed a Targeted proteomics assay which to the best of our knowledge for very first-time enables identification of peptides from 54 meningioma patients via SRM assay to validate the key proteins emerging from our study. This resulted in the identification of peptides from CLIC1, ES8L2, and AHNK many of which are receptors and kinases and are difficult to be characterized using conventional approaches. Furthermore, we were also able to monitor transitions for proteins like NEK9 and CKAP4 which have been reported to be associated with meningioma pathobiology. We believe, this study can aid in designing peptide-based validation assays for meningioma patients as well as IHC studies for clinical applications.

Keywords: label-free proteomics, meningioma tissue, patient-derived primary cell line, RT^2 PCR arrays, Integrin Linked Kinase inhibition, Integrin pathway, PI3K-Akt pathway

INTRODUCTION

Meningiomas are tumors that arise from the outer layering of the brain; the exact site of origin of these tumors is believed to be the arachnoid villi cells. Studies have revealed that meningiomas make up for nearly 35% of all primary brain tumors (1). Although, relatively indolent meningiomas can cause comorbidities including neurological and cognitive disorders. In a study by Stafford et al. via a retrospective study they pointed out that even people with WHO Grade I meningiomas had issues for long term survival owing to impaired cognition (2). With the advent of genomics platforms, the implications of mutations in NF2 as well as Non-NF2 tumors harboring mutations at TRAF7, KLF4, and AKT1 have been recognized (3). Recent research has also highlighted novel insights into the methylation patterns and have proposed the existence of multiple subtypes of meningiomas based on their epigenetic signatures (4). However, not many proteomics studies have been employed for deciphering the perturbations at the protein level.

Though limited in number, there have been recent studies that have uncovered newer insights into the way the tumors behave using several proteomics approaches (5–8). We earlier reported an elevation in the levels of AHNAK, Gelsolin, S100 family of proteins, and several other proteins like CKAP4 that were specific to particular grades of meningiomas (9). Several of these candidates which earlier identified via our quantitative proteomics study were also found to be involved in meningioma pathobiology via the current study and independent studies on larger patient cohorts as well (7, 10).

Therefore, one of the major aims of the current study was to further identify protein markers that can in addition to existing histopathology studies point out cases wherein the patients might need closer follow-up. While contemporary studies have investigated peripheral proteomic alterations of meningiomas; we have employed extensive proteomic and meta-analysis to provide a comprehensive landscape of the protein that has emerged from various studies to provide a panel of proteins that can be screened across a larger cohort of meningiomas via targeted proteomics approach (7, 9). As a first-time attempt, we have developed targeted proteomics-based Selected Reaction Monitoring (SRM) Assays toward clinical translation employing a patient-derived spectral library. This enabled validation of peptides of several candidate biomarkers like VIM, ANXA2, AHNAK, TS101, and CLIC1 from the surgically resected meningioma tissues and control arachnoid regions.

Furthermore, several studies have shown prominent involvement of the Integrin and Focal adhesion pathway in meningiomas including Focal Adhesion Kinase (FAK), ERK. Furthermore, approaches that have inhibited FAK *in vivo* models have shown promising results in terms of inhibition of tumor growth (11). From our current and previous study (9), we have identified major perturbation in components of the Integrin pathway namely ITGAM, ECM2, FBLN1, indicating that there is a strong possibility of involvement of ECM members in meningioma pathobiology. Despite playing a crucial role in recruiting several growth factors as well as engaging in crosstalk with several signaling cascades; Integrins, however, are not known to have any intrinsic enzymatic activity. Many of the signaling aspects of the Focal adhesion junction is modulated via the kinases namely the FAK and the Integrin Linked Kinase (ILK) (12, 13). Thus, inhibition of FAK and ILK has emerged as a promising juncture for therapeutic intervention (14–16). An in-depth mapping of the altered signaling cascades and networks of the identified protein candidates using a plethora of approaches including PTM analysis, machine learning as well as *in silico.* the analysis revealed alterations in Focal adhesion, PI3K-Akt and several components that are downstream effectors of ILK. A previous study using meningioma cell lines reported the efficacy of ILK inhibition in meningiomas (16). To explore whether inhibition of ILK can influence the protein candidates and dysregulated pathways emerging from our study, in this current study we have treated patient-derived primary cell lines with an ILK inhibitor (Cpd22) to assess transcriptional level alterations.

Herein, we present a comprehensive proteomic profile of meningioma patients aimed to look at protein markers for the identification of a panel of protein that can be screened in a large patient cohort using targeted proteomics assay as well as immunohistochemistry. We further provide the first report of SRM assays in meningiomas in nearly 50 patients. Our study also investigates network-level perturbations via machine learning approaches, phosphoproteomics to map perturbed pathways. For the identification of a potential adjuvant therapeutic adjunct in meningiomas where the tumor is either aggressive, recurrent or is unamenable to conventional surgery owing to tumor location; we have assessed the implication of ILK inhibition in meningioma cell lines.

MATERIALS AND METHODS

Sample Collection

All samples were collected following the IEC (Institute Ethics Committee, IIT Bombay) and IRB (Institute Review Board, Tata Memorial Hospital, Mumbai). This study was approved as part of an institutional review board (ACTREC-TMC IEC No.149, Advanced Centre for Treatment, Research and Education in Cancer and Tata Memorial Centre) following approved guidelines. Patients with radiologically suspected meningiomas were enrolled after giving written informed consent. Surgically resected tissues were procured from patients diagnosed to be meningiomas by radiology. Tumor tissues were flash-frozen in liquid nitrogen (**Supplementary Table 1**).

Protein Extraction and Digestion

Global proteomic profiling/shotgun proteomics was performed across MG1 (n = 10), MG II (n = 11), MGIII (n = 2), Normal dura mater (n = 4), and Arachnoid mater (n = 4). Protein extraction was done using Urea buffer (8 M Urea, Tris–HCl buffer) with the addition of Phosphatase inhibitor cocktail (SigmaAldrich®, United States) as mentioned in the protocol by CPTAC Investigators (17). In brief, tumor tissue was washed with 1X PBS and cut (around 75 mg) tissue lysis was performed with sonication followed by bead milling at 90 s for 3 cycles.

The lysates were centrifuged at 12000 r.p.m at 4 degrees for 15 min to clear the debris, the supernatant was quantified using 2D-Quantification kit (Bio-Rad, United States) and 100 μg of protein was digested using Trypsin (Pierce, Thermo Fisher Scientific, United States) for 16 h at 37°C Followed by vacuum drying the peptides and reconstitution with 0.1% Formic Acid. The peptides were quantified using Thermo plate reader using Scope's method 1 μg of the peptide was used for the LC-MS/MS run.

Mass Spectrometry Based Global Proteomic Analysis of Meningioma Patient Cohort and Examination of the Phosphopeptides From the Meningioma Patient Cohort

All patient samples and non-tumor samples were run in the Q-Exactive Orbitrap Mass Spectrometer (Thermo Fisher Scientific, United States) using a gradient of 0.1% FA and Acetonitrile for 240 with blanks after every sample. The scan range was set from 350–1700 m/z and the resolution was set to 17,500. Phosphoproteomic enrichment was done using the TiO_2 enrichment column (Pierce, Thermo Fisher Scientific, United States). In brief protein extraction of surgically resected tissues (MGI, n = 6, MGII, n = 10, and MGIII, n = 2) were done from 18 patient samples, around 300 μg of protein was digested using Trypsin (Pierce, Thermo Fisher Scientific, United States), followed by peptide level enrichment using TiO_2 enrichment column, enriched peptides were further cleaned up using graphite clean-up kit.

Data Acquisition and Analysis by Proteome Discoverer 2.2

All.raw files acquired were analyzed using Proteome Discoverer Version 2.2 using 1% FDR for both the protein as well as peptides. Static modification for the label-free run was set to Carbamidomethyl and for the phosphoproteomics run it was the same with S, T, and Y selected for the variable modification. The database used for protein annotation was Uniprot using Homo sapiens data as background. Additionally, the search engines used were MASCOT and SEQUEST. Total peptide intensity was used for data normalization post missing value imputation on the procured abundances. Metaboanalyst was used for visualization of the heatmaps (18). (**Supplementary Figure 1**: Parameters of analysis in Proteome Discoverer 2.2).

Generation of Transition List for SRM Assay Using Skyline

The transition list of proteins was prepared using the Skyline (Version 4.2) (19). The spectral library generation was done using the raw data as generated from the meningioma patients via the label-free proteomic analysis. The "precursor charges" used were 2, Ion charges used were 1, 2; y ions were selected. Precursor and product masses were set to be "Monoisotopic."

SRM Assay for Differentially Expressed Proteins

The transition list of differentially expressed proteins was prepared using Skyline®. Transitions of specific proteins were optimized along with the LC parameters. All runs were performed using nano-LC mode. The assay was performed in Triple Quadrupole Instrument (TSQ, ThermoFisher Scientific, United States) Peptides (1 μg) was reconstituted in 10 μl 0.1% FA. The peptides were then run in a gradient comprising of solvent A (0.1% FA) Solvent B (80% ACN with 0.1% FA) for 45 min per run. The flow rate was maintained at 300 nl/min and the column used was ES 803 Easy spray pepmap C18 column. The acquisition was done using TSQ Altis™ Triple Quadrupole Mass Spectrometer (Thermo Fisher Scientific, United States) using a method duration of 45 min.

Cell Culture and Inhibition Assay

The primary cells were procured from the Tissue Bank, Imperial College London. The primary cells were seeded from surgically resected tissues of meningioma patients. The cells were passaged prior to freezing in liquid Nitrogen for long term storage. The cells were taken out and revived in DMEM-F12 media. However, owing to the slow growth (doubling time was nearly 50 h) the cells were moved to Waymouth media supplemented with 20% FBS. For the inhibition assay cells were passaged in 6 well plates in triplicates and treated with 2.5 μM of ILK inhibitor namely Cpd22, Merck Millipore. The cells were seeded in two batches "Untreated" (No drug) and "Treated" (With Drug). The treatment was carried out for 24 h. Post-treatment the cells were pelleted down and further taken forward for RNA extraction.

RNA Extraction and cDNA Synthesis

Cells were lysed using RLT buffer and the RNA extraction was done using the Qiagen® RNA Easy extraction Kit (Catalog No: 74104) as per the kit protocol. Briefly, cells were lysed and then mixed with 70% ethanol followed by loading on mini-spin columns. The RNA yield was determined using a spectrophotometer Two-step cDNA synthesis was carried out using the MMLV RT cDNA synthesis kit.

RT^2PCR Using Pathway Specific Arrays

The RT^2 PCR was done using the Qiagen RT^2 PCR array of two pathways namely the Cytoskeletal Regulators and Akt Pathway. The experimental set up included two conditions treated vs untreated cells (Meningioma Primary Cell lines of MGI tumor origin) treated with Cpd22 (Integrin Linked Kinase Inhibitor, Merck Millipore). Mature RNA was isolated using an RNA extraction kit according to the manufacturer's instructions. RNA quality was determined using a spectrophotometer and was reverse transcribed using a cDNA conversion kit. The cDNA was used in the real-time RT^2 Profiler PCR Array (QIAGEN, Cat. no. PAHS-088Z) in combination with RT2 SYBR®Green qPCR Mastermix (Cat. no. 330529). CT values were exported to an Excel file to create a table of CT values. This table was

then uploaded on to the data analysis web portal at http://www.qiagen.com/proteinglobe. Samples were assigned to controls and test groups. CT values were normalized based on the automatic selection from the HKG panel of reference protein (**Supplementary Data 8**).

Pathway Analysis and PPI: Protein-Protein Interaction Analysis

We performed functional annotation clustering analysis for ANOVA passed protein using DAVID (20, 21). Parameters used were custom classification stringency setting; similarity term overlap = 5, similarity threshold = 0.95, initial group membership = 3, final group membership = 3, multiple linkage threshold = 0.5, EASE score = 1.0, and Benjamini-Hochberg correction was applied. Enrichment score was taken into account to narrow down the cluster count. Further, investigation of complex interaction and prediction of pathways among the significant proteins in MG1 vs MG2 were done using Reactome.org (22) and KEGG Database (23). A multi-functional online software NetworkAnalyst 3.0 (24) was used to analyze the protein list for constructing the visualized PPI network. Kinome analysis was performed in kinhub.org with the kinases identified from the dataset. Additionally, REVIGO was used for visualizing the GO terms associated with the phosphopeptides (25). For investigating the effect of Cpd22 on biological pathways of the meningioma cell lines Reactome.org (22) and KEGG (23) were used. To understand the effect of the inhibitor on pathways like PI3K-Akt and Focal Adhesion, the selected pathway entities were extracted from NCI nature (26), KEGG 2019 and Reactome 2019 databases and the list of significant proteins were mapped. Furthermore, the information from the analysis was taken to build a biological pathway model (**Supplementary Data 5, 6, 8, 9**).

MLP and Neural Network-Based Mining for Identification of Key Drivers

Multilayer perceptron (MLP), a 3-layered parsimonious MLP architecture using sigmoid activation functions with a feed-forward – backpropagation learning algorithm was applied to model protein-protein interactions between proteins perturbed in meningiomas as identified via proteomic analysis. ANOVA passed proteins were taken for the Low-Grade vs High-Grade comparison. To define an interaction map for the proteins, the summed weights of the trained ANN model, leading from a given input to a given output, were used to illustrate and score the interaction between proteins (27). The Pearson correlation coefficient r with a cut-off value of 0.7 was implemented in the algorithm to remove the least significant interaction scores. Monte-Carlo cross-validation (MCCV). To prevent the ANN model from being over-trained, an MCCV strategy was applied as follows. The algorithm was coded in C and empirical work on the dataset was presented in the subsequent section. Visualization of interactome network maps The Cytoscape software platform (version 2.8) for molecular interaction display was used in this study. The above approaches were adapted and modified for the current data set from (27–29) (**Supplementary Data 4B**).

RESULTS

Clinical Parameters

The patients were examined via MRI for determining the tumor location. Additionally, surgically resected tissues were examined by standard histopathology as per the WHO guidelines for assigning the grades. For the Global proteomics study, 10 MGI (Benign) and 11 MGII (Atypical) cases were taken forward (**Supplementary Data 1**).

Label-Free Quantification of Surgically Resected Meningiomas to Identify Altered Proteomic Signatures

Patients with radiologically suspected meningiomas were accrued for the study. Patients were assigned WHO Grades namely MGI, MGII, and MGIII post histopathological examination. The information of the accrued patients is mentioned in **Supplementary Table 1**. Individual proteomic analysis of the patients was performed in LC-MS in Orbitrap Q-Exactive which enabled the identification of nearly 5659 Master Proteins with 1% False Discovery Rate. Further filtering based on $\geq$2 unique peptides enabled the identification of nearly 4600 proteins with $\geq$2 unique peptides across the cohort with high confidence. Further, Log_2 transformation of the obtained abundances was performed which were taken forward for further statistical analysis (**Figure 1**, **Table 1**, and **Supplementary Data 2**, and **Supplementary Figure 1**).

Evaluating Grade-Wise Markers in Meningioma Patients, Correlation With Histopathological Features

The comparison of the significant proteins as obtained post ANOVA test was further subjected to Unsupervised clustering which revealed that there were three clusters (**Figure 2A**). As per the heatmap it is visible that NDRG1, MRC2, and COPS8 is highly abundant in the higher grades of meningiomas whereas more abundances of COL14A, COL12A1 were found in the non-tumor controls (**Figures 2A,B**). Furthermore, a volcano plot analysis of the entire dataset with MG vs Non-tumor controls revealed significant upheaval in proteins like SPAG9, NDRG2 (**Figure 2D**). Comparison of MGI (n = 10) and MGII (n = 11) enabled the identification of nearly 230 proteins with high confidence that were significantly ($p \leq 0.05$) altered in MGI vs MGII. NBAS, TIMP2, and NCK1 were few of the key proteins found to be altered in Atypical meningiomas (**Figures 2F,G** and **Supplementary Data 3**). Furthermore, correlation of the proteomic and histopathological subgroups was found out in the patient cohort with WHO classified patients that were annotated MGI; however, had bone invasion clustered with the Atypical patients indicating the possibility of intratumor heterogeneity and a need for closer follow up of these patients (**Figure 2A**). We probed the 672 proteins that have significantly passed ANOVA for the Gene Ontology. Grade-wise comparisons on MGI (benign) vs MGII (atypical) meningiomas revealed prominent involvement of TIMM50, SEC23A, MTAP, RPS23,

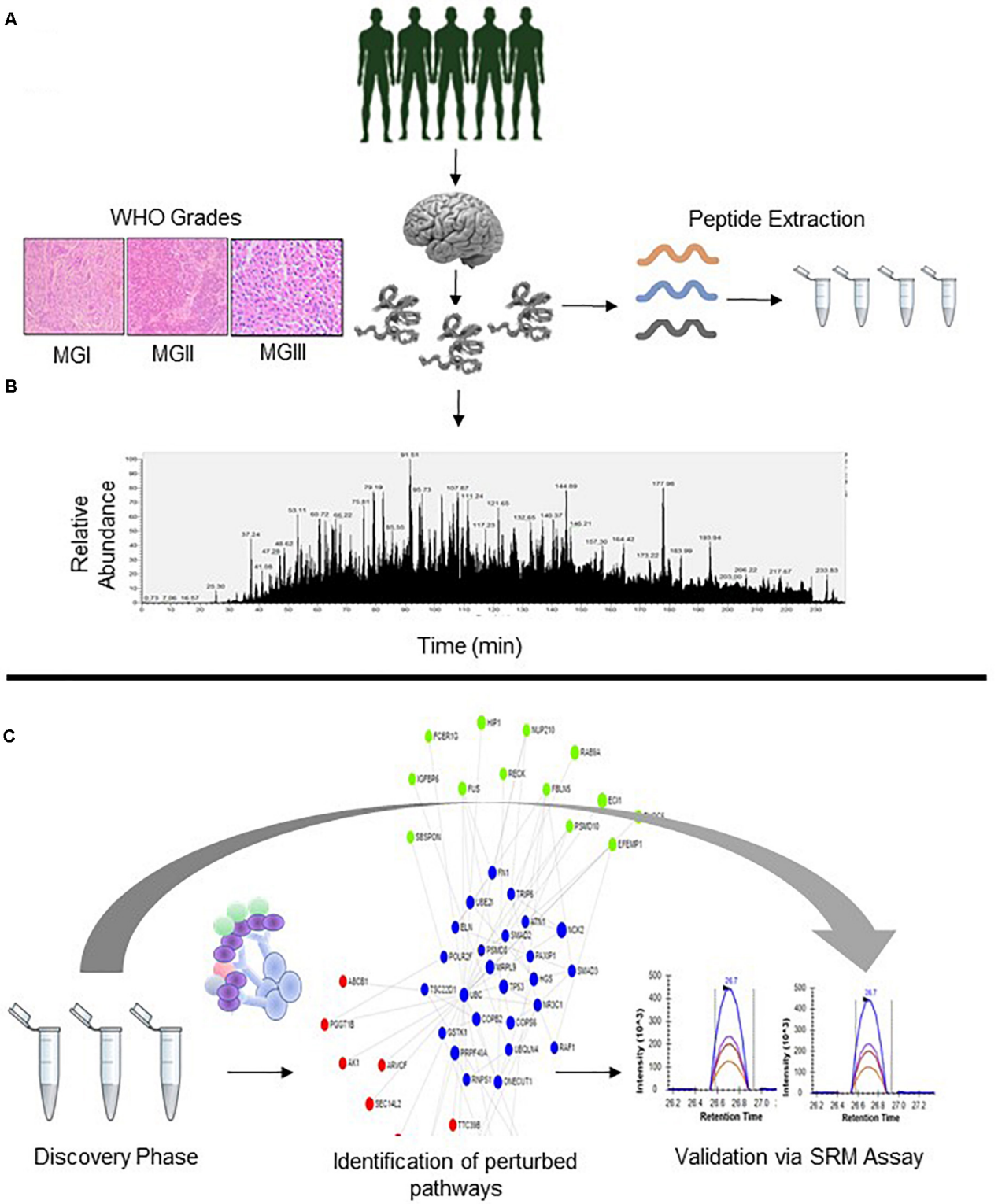


FIGURE 1 | Overall Schematic of Proteomic Analysis of Meningioma Patients. **(A)** Patients suspected to have meningioma **(B)** Analysis of WHO Grade-specific proteomic alterations; patient segregation based on WHO classification **(C)** Gene Ontology, PPI, and Pathway mapping and development of SRM Assays for the analyzed proteins.

TABLE 1 | Key candidate proteins with role in meningioma pathobiology, monitored peptides via SRM Assay.

Protein	Avg. Abundance (Log Scale)			Peptides Monitored via SRM Assay	Role in meningiomas/ Reported via other studies	Phosphosite identified	Influence on signaling pathway(s)	Inhibitors/Therapeutic intervention points
	NT (*n* = 8)	MGI (*n* = 10)	MGII (*n* = 11)					
VIM (P08670)	34.7	36.0	35.2	R.DNLAEDIMR.L K.LQEEMLQR.E R.QDVDNASLAR.L	Phosphorylated form of Vimentin ascribed in non-infiltrative meningiomas (39, 40)	Phospho [S7; T20; S22; S25; S26; S34; S39; S51; Y53; S55; S56; S72; S73; S83; S87; S214; S226; S325; S339; S409; S412; S419; S420; T426; S430; S438; T458; S459]	Vimentin a component of extracellular matrix also is a substrate for multiple kinases (41)	Withaferin-A (WFA) (41)
ANXA2 (P07355)	31.3	33.1	32.8	K.AYTNFDAER.D R.QIAFAYOR.R K.TPAQYDASELK.A	Annexin A2 as well as key interactor S100A10 has been reported to be elevated in meningiomas (42–44)	Phospho [S12; Y24; S26; S89; S92; Y188; Y199]	Annexin A2 is a multifunctional protein mediating actin remodeling, membrane assembly, signal transduction, etc. (45, 46)	Tri-Substituted 1,2,4-triazoles inhibitANXA2-S100A10 interaction (42)
CKAP4 (Q07065)	26.6	28.5	28.7	R.LPPQDFLDR.L K.ASVSQVEADLK.M R.TAVDSLVAYSVK.I	CKAP4 has been seen to be aberrantly expressed in meningiomas via multipronged approaches including IHC (7)	Phospho [S460; S461]	CKAP4 mediates activation of the PI3K-Akt pathway.	
CLIC1 (O00299)	25.7	27.5	27.3	K.IGNCPFSQR.L K.LAALNPESNTAGLDIFAK.F R.GFTIPEAFR.G	Reported to be upregulated in meningiomas especially in atypical and anaplastic types (7)	NA	Regulates MAPK/Akt pathways (47)	Biguanide-related drugs (47)
EIF4G1 (Q04637)	24.2	25.7	26.7	K.QVTVLAIDTEER.L K.GVIDLIFEK.A K.VEYTLGEESEAPGQR.A	Critical component of translational initiation. Known to function as an oncoprotein in several cancers (44)	Phospho [S1185; S1187]	Translational machinima as well as influence PI3K/AKT/mTOR pathway	Silvestrol has ha inhibiory effects on EIF family in meningioma cell lines (48)

Details mentioned in ***Supplementary Data 2, 6, 9A****.*

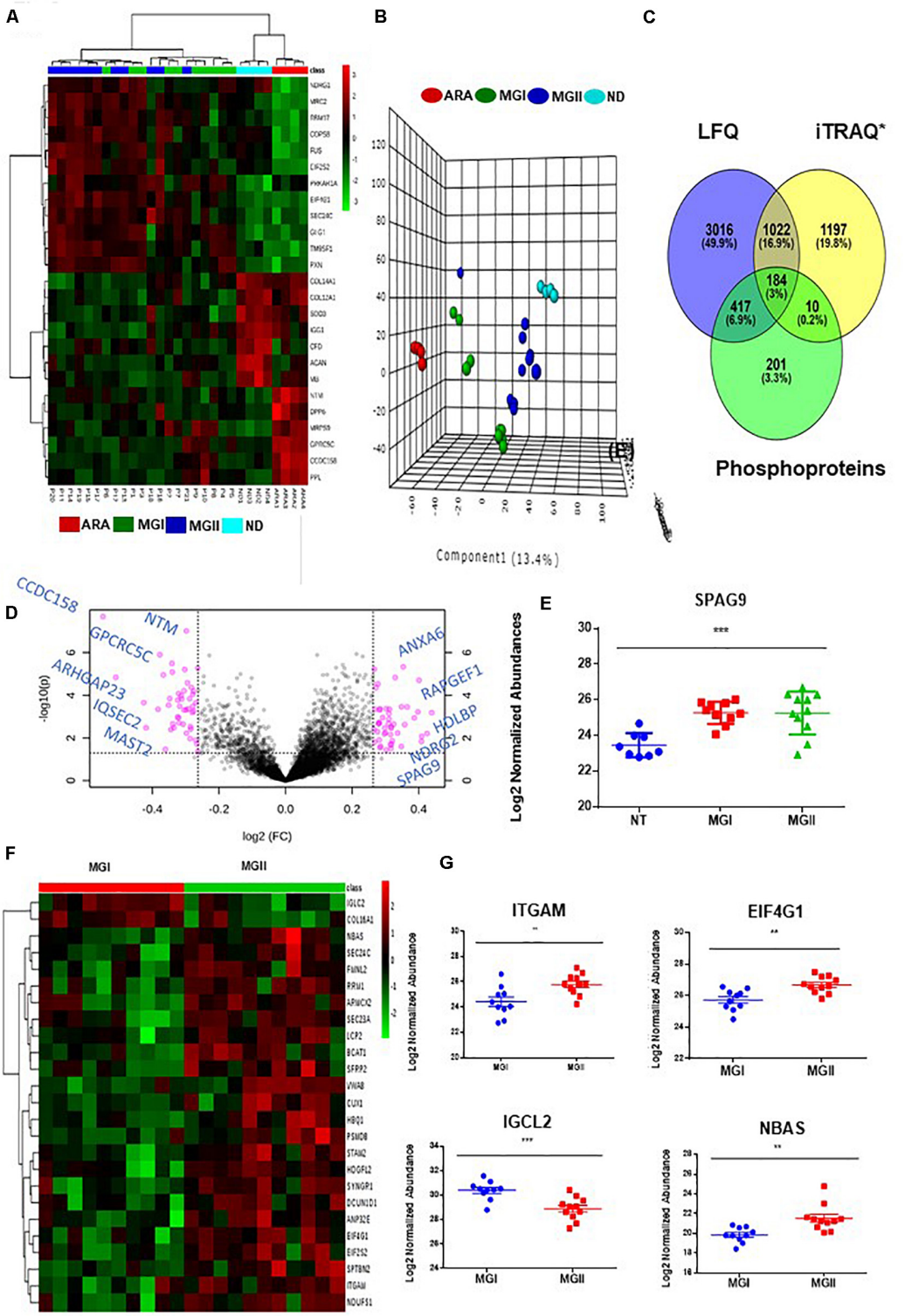


FIGURE 2 | Identification of meningioma subtypes using proteomics and examining the perturbed networks. **(A)** Unsupervised clustering (Distance measure: Euclidian; Clustering algorithm: Ward) reveals distinct subgroups meningiomas **(B)**. Principle Component Analysis reveals the segregation of MGI, MGII, Dura, and Arachnoid; few MGI cases seemed to exhibit proteomic profiles as that of the atypical tumors **(C)** Cross comparison of Meningioma Label-free proteomic data set from individual patients with quantitative grade-wise proteomics study and PTM analysis. **(D)** Volcano plot (MG vs Non-tumor controls) Log FC $\geq$ 1.2. **(E)** Log 2 Normalized abundance across grades and non-tumor control of SPAG9 **(F)** Supervised Hierchial Clustering reveals grade-specific protein profile in MGI vs MGII cases **(G)** Key altered proteins between MGI vs MGII (*t*-test passed; $p < 0.05$).

and ADRM1 based on the concordance analysis of highly ranked features from across 20 Machine Learning randomized iterations obtained via the methods of the Ball group (28) (**Figures 3A,B** and **Supplementary Data 4B, 5**).

The functional annotation clustering based on Molecular Function; Biological Process, and Cellular component provides 135 clusters taking the 672 proteins into account which were significant in meningiomas vs non-tumor controls. The top hits in terms of Biological Processes were, cell-cell adhesion (p-value = 1.39E-13, ES = 14.61) and Extracellular Matrix Organization (p-value = 1.53E-06, ES = 4.664). In the ECM clusters, a hub of proteins that were exclusively found to be related to collagen biosynthesis was COL4A4, COL6A2, COL14A1, COL16A1, and COL18A. Several members of extracellular matrix remodeling and Integrin signaling like ECM2, LUM, ITGB4, ITGM, and Fibulin family were also mapped. The proteins of these clusters were related to the PI3K-Akt signaling pathway and Focal adhesion. On the other hand, in cell-cell adhesion Annexin A1, Annexin A2, the hub of the actin-binding protein, and proteins related to translational machinery were also found (**Supplementary Data 4A**). The Gene Set Enrichment analysis of overall identified significant proteins revealed the involvement of Focal Adhesion, PI3K-Akt, Hippo, VEGF, and Ras signaling pathways with possibilities of crosstalk among the dysregulated pathways. **Figures 3A,B**, **4A** (**Supplementary Data 4A**).

Identification of Phosphorylated Proteins From Meningioma Patients

Investigating the global proteome level alterations also lead to the identification of proteins that were involved in key signaling cascades. To assess the post-translational modification of these proteins we investigated the phosphorylation status in the patient group. Phosphoproteomics was performed using TiO_2 based enrichment approach on 18 patients to identify the phosphorylation status of the S, T, and Y of the key proteins that were found to be altered in meningioma patients. We were able to identify nearly 812 proteins that had $\geq$2 unique peptides with high confidence. This study revealed the identification of phosphorylated sites of several proteins that were found to play a role in meningioma pathobiology by several studies including EPB41L2, NDRG2, SPTB, MAGED2, MXRA7, and nearly 51 Kinases which were also mapped further through Kinome Analysis (**Table 2**, **Supplementary Data 6** and **Figures 2C**, **5A–C**).

Collation of Altered Proteomic Signatures Procured From Individual Patients With Grade-Wise Differential Surrogate Markers

We had previously identified several differentially expressed proteins in various grades of meningiomas using iTRAQ based proteomics study (9); 1206 of the differentially expressed proteins were also found out in the individual label-free proteomic analysis substantiating the prominent roles of these in meningioma pathobiology. Some of the key proteins identified via both the approaches include several regulators and cytoskeletal components like AHNAK, PLEC, FLNB, and DSP; routinely used IHC markers like VIM, S100 and Annexin family proteins were also identified with high confidence (1% FDR, $\geq$2 unique peptides; **Supplementary Data 7**).

Perturbation of Several Integrin Components and Influence of Inhibition of Integrin Linked Kinase in Meningioma Primary Cell Line

The global proteomic analysis of meningioma patients in the current study revealed perturbations in several components of Integrin including ITGAV, ITGB2, ITGA2B, and ILKAP (**Supplementary Table 2** and **Supplementary Figure 6**). Furthermore, comparative analysis of Integrin components with Sharma et al., 2015 (9) revealed perturbations in ILKAP, ITGA6 as well as a kinase named ILK which is a key mediator of the integrin pathway and several downstream cascades. Further *in silico* analysis enabled the identification of major interacting nodes in the Integrin and PI3K-Akt pathways which could be downstream effectors of ILK (**Supplementary Data 4A**).

The inhibition of primary cell lines of meningioma with ILK inhibitor (Cpd22) was performed in meningioma patient-derived primary cell lines. The transcriptome level alterations in various components of PI3K-Akt and Cytoskeletal modulators were identified via transcriptomics level analysis using pathways specific RT^2 PCR arrays. Key candidate genes altered in the Control vs Treated were MTOR, RHOA, EIF4G1 from the PI3K-Akt pathway and MSN, CYFIP2, AURKC in the Integrin pathway (**Figures 4B,C**, **Supplementary Table 8**, and **Supplementary Figures 2–4**).

Corroboration of Altered Protein Signatures to Evaluate Relative Abundance Across Individual Patient Cohort via SRM Assay Employing a Patient-Specific Spectral Library

Using the individual patient spectra, we prepared a patient-specific "Spectral Library" for developing a targeted proteomics assay to validate key proteins. We monitored transitions for proteins like VIM, ANXA2, S100 family proteins. Also, we validated targets like AHNAK, tumor-specific proteins like TPD52L2, TS101. Peptides corresponding to NEK9, SELENBP1 which were previously reported as a marker for aggressivity via two independent proteomics study were also optimized and monitored across the meningioma patient cohort (5, 7). We further monitored the levels of PI3K-Akt and Integrin family members like EIF4G1 and CKAP4 (**Figure 6**, **Supplementary Data 7, 9**, and **Supplementary Figures 5, 6**).

DISCUSSION

Our study employs a plethora of approaches to look into pathway-specific perturbations in meningioma tissue with

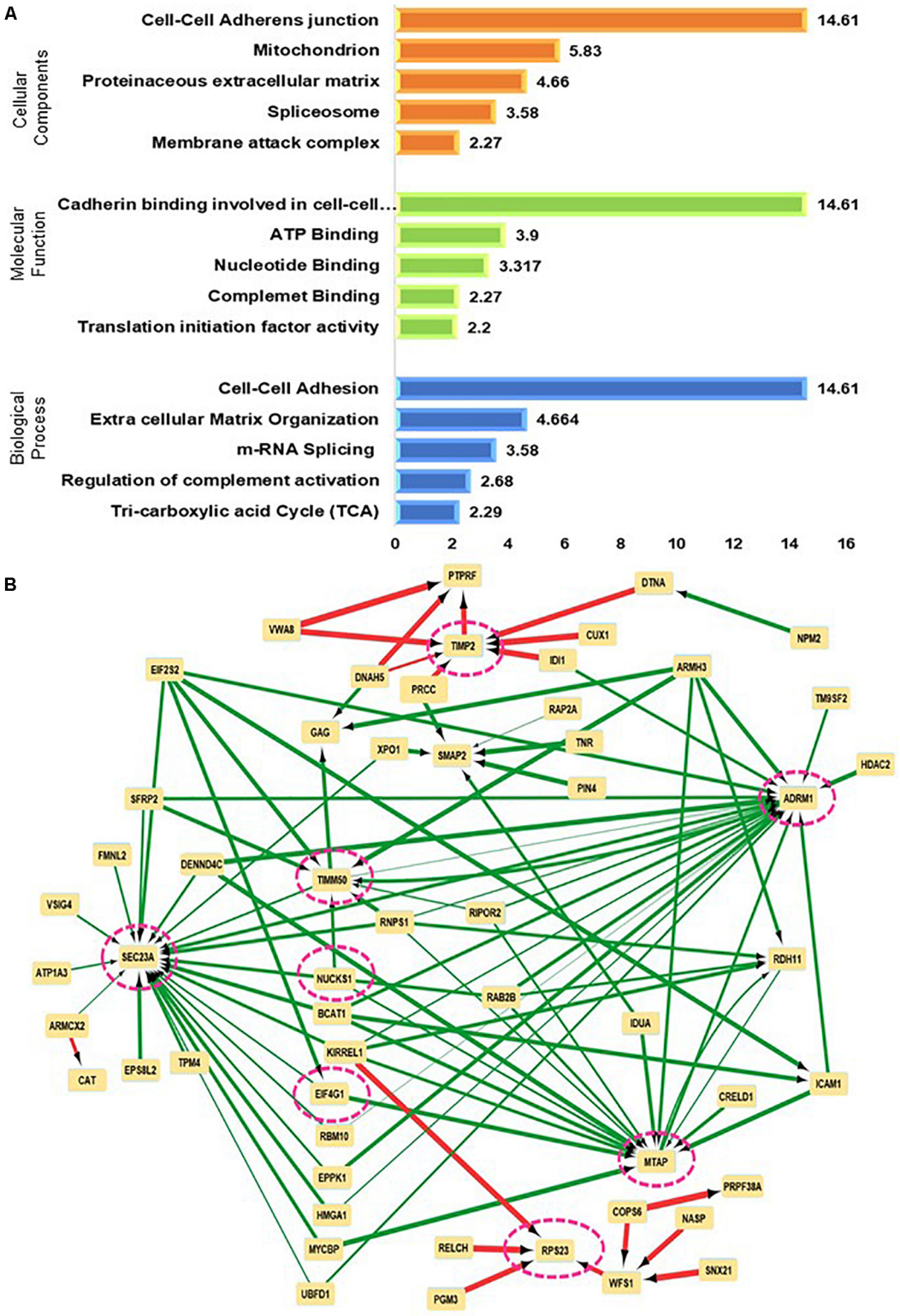


FIGURE 3 | Mapping perturbed pathways and networks from grade-wise meningioma tissue comparisons **(A)** Functional enrichment test on significant proteins emerging from comparisons of different grades of meningioma **(B)** Neural networking of significantly altered proteins using lower grade vs higher-grade meningiomas as predicted via a Multilayer perceptron (MLP) neural network architecture.

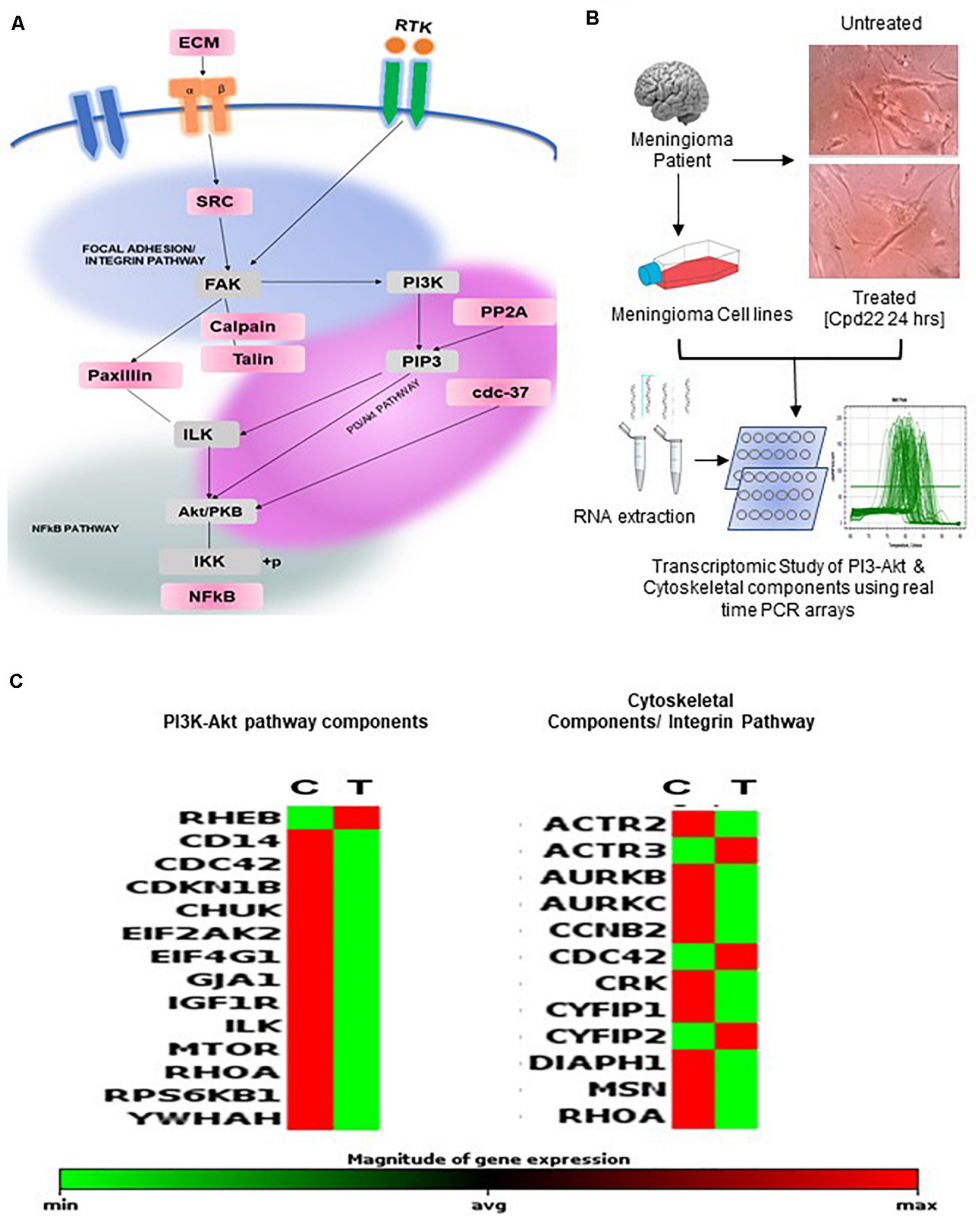


FIGURE 4 | Assessment of perturbations of Integrin and PI3K-Akt pathway in meningiomas: **(A)** Hypothesis of possible crosstalk among PI3-Akt, NFkappaB, and Focal adhesion pathways **(B)** Workflow depicting Patient-derived primary cell line which was treated with 2.5 μM of ILK inhibitor and examined for alterations in the transcriptome with a specific focus on PI3-Akt pathway and Cytoskeletal Regulators **(C)** Alterations of PI3K-Akt and Cytoskeletal components as derived from the via transcriptomic analysis.

an emphasis on the Integrin and PI3K-Akt components employing targeted transcriptomics as well as proteomics approaches. To the best of our knowledge, this is the first study that employs targeted proteomics workflow for monitoring peptides from surgically resected tissues in meningioma patients. The protein markers that have emerged from several studies were screened using this approach and thus can be further used for designing peptide-based assays and immunohistochemistry (IHC) based studies for meningioma patients (7, 9, 10).

Potential of the Label-Free Proteomic Analysis on Identifying Protein Markers and Use of Patient-Derived Spectral Libraries

Meningiomas are one of the most prevalent primary tumors originating from the outer layer of the brain. Several recent studies have highlighted the novel features of these tumors, including alterations at the genetic, epigenetic, and proteomic levels, which has potentiated the possibility of

TABLE 2 | Key proteins and Kinases Identified via the phosphoproteomic analysis followed by Kinome Mapping and their potential role in meningioma pathobiology (**Supplementary Data Table 6**).

Protein	Identified Phosphosite	Potential role in meningiomas	Influence on signaling pathway(s)	Points of therapeutic intervention
Q9H1E3 (NUCKS1)	Phospho [S19; S54; S58; S61; S181; T202; S204; S214; T220; S221; T222; S223]	Modulates chromatin structure and key events of replication and transcription. It is known to be phosphorylated by CK2, Cdks, and is increasingly being noted as a cancer biomarker (49). Reported to be upregulated in atypical and anaplastic meningiomas (5)	Influences Insulin mediated signaling, NFkB as well as Myc (49)	miR-142-3p has been reported to inhibit NUCKS1 (50)
Q02952 (AKAP12)	Phospho [S219; S280; S283; T285; S286; S312; S371; T597; S598; S627; S629; S696; S697; S698; S1587]	AKAP12 has been reported as a central regulator of meningioma aggressiveness with a possible role in progression (5).	Known to downplay events linked with metastasis and invasiveness. It acts as an antagonist to VEGF (51)	A few peptide inhibitors and a small molecule inhibitor FMP−API−1 have shown to disrupt the AKAP-PKA (52)
Q9P0K1 (ADAM22)	Phospho [S832; S834; S857]	The ADAM22 gene encodes a protein that interacts with LGI1 and forms a complex. It is known to interact with integrins in the brain as well (53)	ADAM22 is known to interact with NRG1 which in turn phosphorylates AKT (53).	Pidolic acid (53)
P08648 (ITGA5)	Phospho [S127; S128]	Integrin alpha-5/beta-1 (ITGA5:ITGB1) is a receptor for fibronectin and fibrinogen (54).	Associated with cell adhesion, focal adhesion functionalities (55).	Resveratrol (53)
Q96B36 (AKT1S1)	Phospho [S88; S92; S202; S203]	AKT1S1 is a proline-rich substrate of AKT that binds to 14-3-3 protein when phosphorylated	Play a crucial role in mediating mTOR and PI3K-Akt signaling cascades	OSI-027 (56)
Q92597 (NDRG1)	Phospho [S326; T328; S330; S332; S333; T335; S336; S364; T366; S367; T375; T699; T699]	Protein encoded by NDRG1 is involved in p53-mediated caspase activation and apoptosis (54). NDRG2 another family member of the NDRGs is known to play role in meningioma tumor recurrence (54)	Involved in NF-kappa B Signaling PI3K/Akt Signaling (54)	Calcium, Okadaic acid (54)
Q9UN36 (NDRG2)	Phospho [S328; T330; S332; S338]	NDRG2 is thought to act as a tumor suppressor. Loss of NDRG2 and decreased presence at genetic level detected in higher grades of meningiomas (54).	Regulates cyclin D1, exerts anti-tumor effects via activation of p38 MAPK. Modulates cell survival and invasion (54, 57).	Fostamatinib (54)

identifying markers that can be used in the clinics for better prognostication of patients (7, 30, 31). There is an imminent need for markers that can be used for prognosis as well as potential targets for therapeutic interventions. The latter is particularly important in cases where the tumors might not be amenable for complete surgical resection or in those cases where a lower grade tumor might have the potential to turn aggressive. In our current study, we have employed label-free proteomics to decipher proteomic alterations in meningioma patients. Using a label-free proteomics approach we identified 7978 Protein groups; 52695 Peptide groups and 1043972 PSMs. The study enabled the identification of nearly 4600 proteins with 1% FDR (≥2 Unique peptides) from surgically resected tumor tissues. Additionally, we have used a "patient-derived spectral library" generated via the global proteomic analysis to develop a targeted proteomic assay workflow for validating several of the altered proteins.

Protein Markers That Can Aid in the Prediction of Patient Prognosis

We compared the proteomic landscape emerging from the various meningiomas stratified based on the WHO guidelines. The study revealed the identification of 672 proteins that were significantly altered among the meningioma grades and non-tumor controls namely the arachnoid and dura regions and 235 proteins were found to be dysregulated among the MGI and MGII patient cohort indicating the possibility of these proteins to be grades-specific. Additionally, using "unsupervised clustering" we observed that a few cases of even lower grades did segregate with atypical cases. On closer examinations of their histopathological signatures, it randomized that some of the patients had bone invasion. The elevated abundance of proteins like NDRG1, MRC2, and FUS which were found to be higher in the atypical cases in a few of the MGI cases indicates a possibility that these cases might have a chance of recurrence in future and warrants a closer

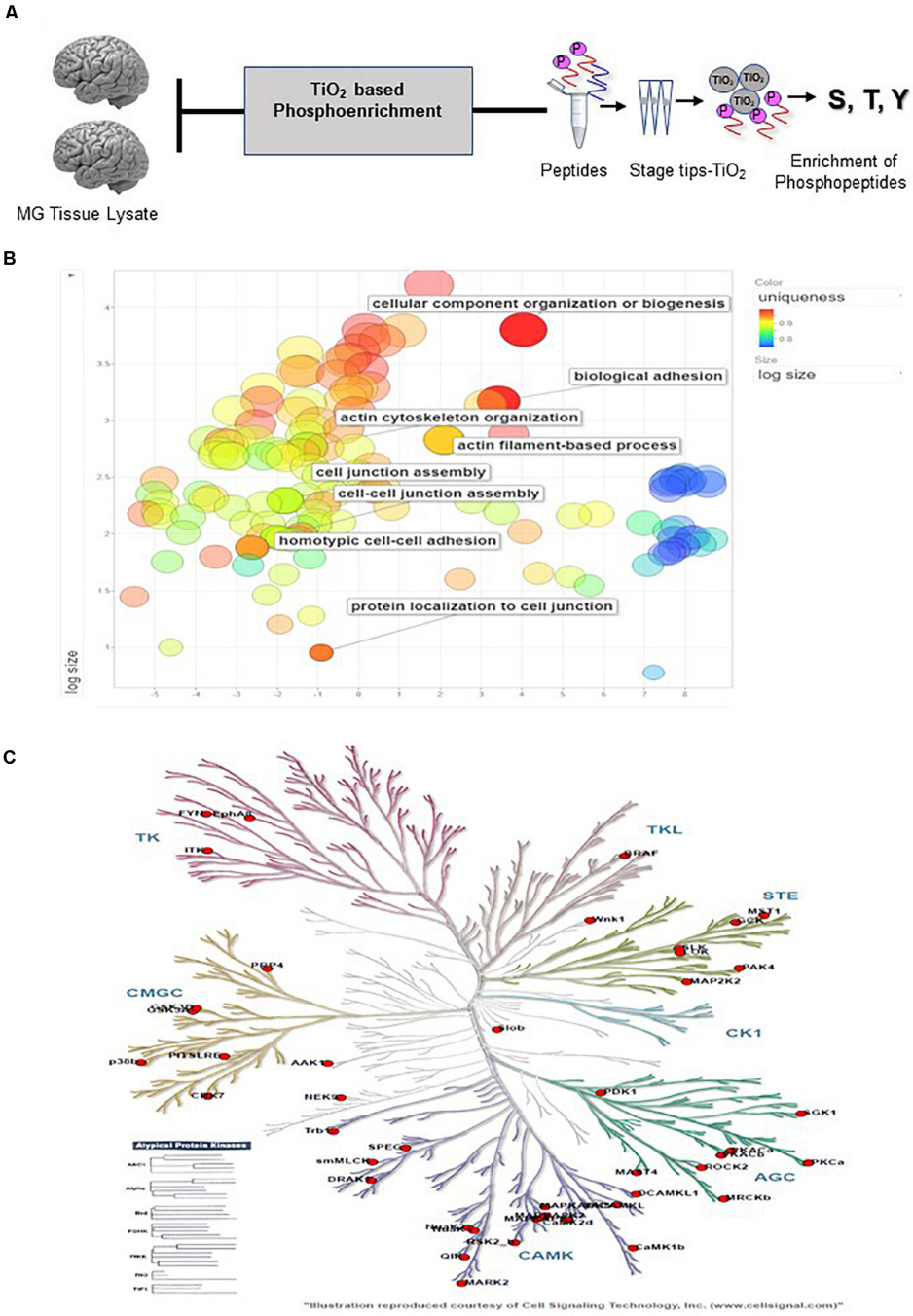


FIGURE 5 | Outcomes of PTM Analysis **(A)** Workflow for enriching the samples for Phosphopeptide enrichment using Titanium dioxide Stage. **(B)** REVIGO Analysis revealed the Gene Ontology-based enrichment of the identified phosphopeptides **(C)** Kinome Analysis performed in http://kinhub.org enabled identification of the Kinome map of the kinases identified via study (**Supplementary Data 6**).

follow up. It is interesting to note that NDRG1 is a known tumor suppressor that has been reported to be associated with stress and the prevalence of hypoxic conditions in tumors. Increased expression of NDRG1 is associated with poor patient outcomes which might be the case for meningiomas as well (32, 33).

Components of Integrin and PI3K-Akt Pathways That Are Involved in Meningioma Pathobiology

Our study points out the prominent involvement of several components of the Focal Adhesion and Integrin pathway in

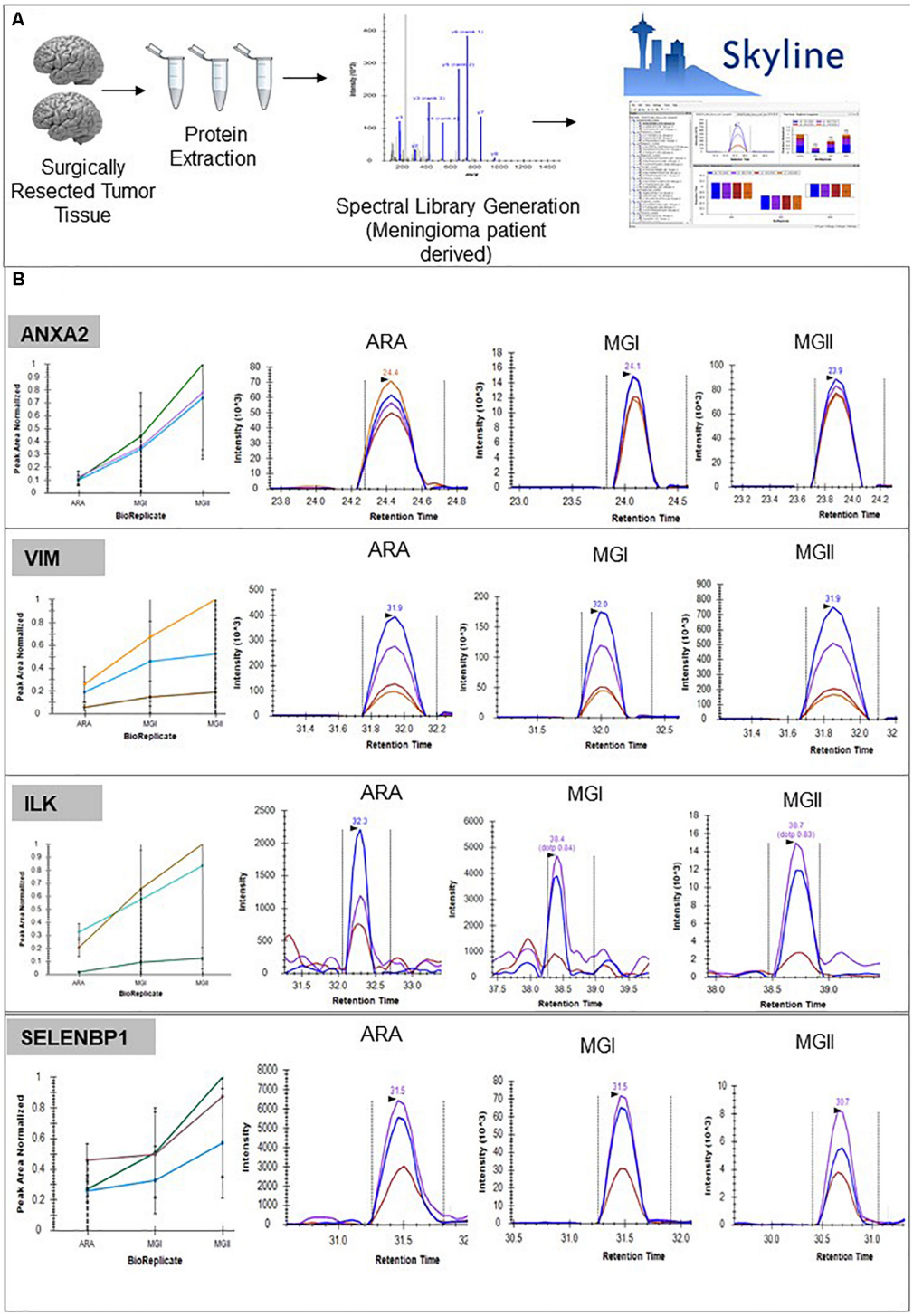


FIGURE 6 | Panel of proteins identified as potential biomarkers. **(A)** Pipeline followed for developing peptide level validation of protein markers from meningioma tissue using targeted proteomics approach including using the global proteome data as a spectral library and using Skyline® for further assay development **(B)** Validation of biomarkers using SRM-based assays. (Details of peptides screened, Peak Areas detailed in **Supplementary Data 8**, **Supplementary Figure 5**).

meningiomas which has also been reported by other studies (34, 35). We found higher levels of FMNL2, ITGAM, and MRC2 among several other components mapping to the Focal adhesion and PI3K-Akt pathway in our patient cohort. While several candidates of the Focal Adhesion family have been associated with meningiomas via transcriptomic and proteomics studies;

it is to be noted that most Integrins by themselves do not possess enzymatic activity (36). ILK, a 59 kDa protein kinase has emerged as an interesting player in mediating the signaling cascades through its interaction with members of the Integrin family as well as engaging several downstream effectors including AKT, GSK3, mTOR (37).

Scope of ILK Inhibition as an Adjuvant Therapy

A study employing the inhibition of ILK showed that this affects cell growth in meningiomas by hindering the interaction of ILK with downstream components of PI3K-AKT (16). To determine the influence on inhibition of the ILK, one of the key components that regulate the three perturbed pathways namely Integrin, PI3K-Akt, and NF-κB, we have used the ILK inhibitor Cpd22 (Merck Millipore) on both meningioma patient-derived primary cell line as well as Ben-Men1. Inhibiting the activity of ILK affected the levels of key proteins known to be key players in meningioma pathobiology namely EIF4G1, CSNK2A1, and several others. At the transcriptomic level we observed the downregulation of EZR, GSN, IQGAP1, IQGAP2, and the several other Cytoskeletal components; AKT1, EIF4G1, ILK, and NFKB1 associated with PI3K-Akt and NF-κB. These findings need to be extrapolated in animal models for further substantiation of these observations. Overall, the perturbations observed using Cpd22 indicate that ILK inhibition has the potential to be used as an adjuvant therapy especially in cases wherein the tumor location makes complete resection difficult as well as for cases of multiple incidences of recurrence.

Designing SRM Assay to Validate Clinically Relevant Protein Markers From Meningioma Patients

In a recent study by Diamandis et al., a global proteomic analysis of meningioma patient cohort of nearly 60 patient FFPE blocks were analyzed and many of the biomarkers that were validated in our current study namely SELENBP1, NDRG1, and MAPK3 were found aberrated in the FFPE samples indicating their true potential as markers for meningiomas that can even be used in clinics (10). Hence, for the first time SRM assay using a "patient-derived spectral library" was employed to monitor several potential meningioma biomarkers. Key candidates of the Focal adhesion pathway like ILK, EZR, MAP4, and VIM was also monitored from a larger patient cohort. The targeted proteomics approach enabled monitoring the levels of the several putative markers namely CLIC1, ES8L2, AHNAK many of which are receptors, and kinases and are difficult to be characterized using conventional approaches. Furthermore, using both label-free based proteomics as well as targeted proteomics we were able to cross-validate many of the proteins namely ANXA2, AHNAK, and CKAP4 that were identified in our earlier study (9).

CONCLUSION

Combining global and targeted patient-derived data is a novel approach and it has the potential to provide cues regarding the meningioma pathobiology which can be useful in longitudinal studies for monitoring the patient prognosis. While inhibitory studies against FAK, mTOR, and Akt components in meningioma patients are ongoing (38); our study provides mechanistic insights on how components of Cytoskeletal regulators, PI3K-Akt, mTOR, and NFκB can be targeted using a single molecule namely the Cpd22 which was able to downplay these components in meningioma primaries. We also lay the foundation for the use of targeted proteomics for validation of key proteins directly from the patient cohort which if integrated with existing modalities of diagnosis and treatment can aid in enhanced patient management.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics datasets have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD014852, PXD014853, PXD014855, and PXD014823.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by TMH and IITB IEC. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SM carried out all experiments and compiled the manuscript. SS, AMo, SE, and SM conceived the study and secured funding. AMo, PS, and SE performed the clinical, surgical, and pathological examination of the patients. DB performed GSEA analysis. GR and PJ performed machine learning-based analysis. SM performed SRM assay experiments. RG was involved in optimizations. NS facilitated MG cell line assays. RT^2 PCR assay was performed by JS and QZ. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded through MHRD-UAY Project UCHHATAR AVISHKAR YOJANA; UAY-(MHRD), project #34_IITB (2016) to SS, MASSFIIT (Mass Spectrometry Facility, IIT Bombay; BT/PR13114/INF/22/206/2015) for MS-based proteomics work, DBT India and The Newton-Bhabha Fellowship (awarded to SM).

ACKNOWLEDGMENTS

Sincere thanks to the MASSFIIT (Mass Spectrometry Facility, IIT Bombay) for the technical support, NIMHANS Brain Bank facility for providing Normal brain tissues. We are grateful to Dr. Vijaya Ramesh, Center for Genomic Medicine, Harvard Medical School for kindly gifting the Ben-Men1 cell line.

Note of gratitude to Sujata Walunj, Sohini Saha, for their involvement in initial optimizations, Saicharan Ghantasala, Sandip Kumar Patel, and Vipin Kumar for technical assistance in TSQ Triple Quadrupole Instrument, Q-Exactive LC-MS/MS, and Orbitrap Fusion, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01600/full#supplementary-material

REFERENCES

1. Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro Oncol.* (2013) 15(Suppl. 2):ii1–56. doi: 10.1093/neuonc/not151
2. Van Alkemade H, de Leau M, Dieleman EMT, Kardaun JWPF, Van Os R, Vandertop WP, et al. Impaired survival and long-term neurological problems in benign meningioma. *Neuro Oncol.* (2012) 14:658–66. doi: 10.1093/neuonc/nos013
3. Clark VE, Erson-Omay EZ, Serin A, Yin J, Cotney J, Ozduman K, et al. Genomic analysis of non-NF2 meningiomas reveals mutations in TRAF7, KLF4, AKT1, and SMO. *Science.* (2013) 339:1077–80. doi: 10.1126/science.1233009
4. Paramasivam N, Hübschmann D, Toprak UH, Ishaque N, Neidert M, Schrimpf D, et al. Mutational patterns and regulatory networks in epigenetic subgroups of meningioma. *Acta Neuropathol.* (2019) 138:295–308. doi: 10.1007/s00401-019-02008-w
5. Parada CA, Osbun J, Kaur S, Yakkioui Y, Shi M, Pan C, et al. Kinome and phosphoproteome of high-grade meningiomas reveal AKAP12 as a central regulator of aggressiveness and its possible role in progression. *Sci Rep.* (2018) 8:2098. doi: 10.1038/s41598-018-19308-y
6. Sharma S, Ray S, Moiyadi A, Sridhar E, Srivastava S. Quantitative proteomic analysis of meningiomas for the identification of surrogate protein markers. *Sci Rep.* (2014) 4:7140. doi: 10.1038/srep07140
7. Dunn J, Ferluga S, Sharma V, Futschik M, Hilton DA, Adams CL, et al. Proteomic analysis discovers the differential expression of novel proteins and phosphoproteins in meningioma including NEK9, HK2 and SET and deregulation of RNA metabolism. *EBioMedicine.* (2019) 40:77–91. doi: 10.1016/j.ebiom.2018.12.048
8. Barkhoudarian G, Whitelegge JP, Kelly DF, Simonian M. Proteomics analysis of brain meningiomas in pursuit of novel biomarkers of the aggressive behavior. *J Proteomics Bioinform.* (2016) 9:53–7. doi: 10.4172/jpb.1000389
9. Sharma S, Ray S, Mukherjee S, Moiyadi A, Sridhar E, Srivastava S. Multipronged quantitative proteomic analyses indicate modulation of various signal transduction pathways in human meningiomas. *Proteomics.* (2015) 15:394–407. doi: 10.1002/pmic.201400328
10. Papaioannou M-D, Djuric U, Kao J, Karimi S, Zadeh G, Aldape K, et al. Proteomic analysis of meningiomas reveals clinically-distinct molecular patterns. *Neuro Oncol.* (2019) 21:1028–38. doi: 10.1093/neuonc/noz084
11. Gogineni VR, Nalla AK, Gupta R, Gujrati M, Klopfenstein JD, Mohanam S, et al. α3β1 integrin promotes radiation-induced migration of meningioma cells. *Int J Oncol.* (2011) 38:1615–24. doi: 10.3892/ijo.2011.987
12. Widmaier M, Rognoni E, Radovanac K, Azimifar SB, Fassler R. Integrin-linked kinase at a glance. *J Cell Sci.* (2012) 125:1839–43. doi: 10.1242/jcs.093864
13. Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer.* (2010) 10:9–22. doi: 10.1038/nrc2748
14. McDonald PC, Fielding AB, Dedhar S. Integrin-linked kinase–essential roles in physiology and cancer biology. *J Cell Sci.* (2008) 121:3121–32. doi: 10.1242/jcs.017996
15. Tiede S, Meyer-Schaller N, Kalathur RKR, Ivanek R, Fagiani E, Schmassmann P, et al. The FAK inhibitor BI 853520 exerts anti-tumor effects in breast cancer. *Oncogenesis.* (2018) 7:73. doi: 10.1038/s41389-018-0083-1
16. Mercado-Pimentel ME, Igarashi S, Dunn AM, Behbahani M, Miller C, Read CM, et al. The novel small molecule inhibitor, OSU-T315, Suppresses vestibular schwannoma and meningioma growth by inhibiting PDK2 function in the AKT pathway activation. *Austin J Med Oncol.* (2016) 3:1025.
17. Zhang H, Liu T, Zhang Z, Payne SH, Zhang BB, McDermott JE, et al. Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell.* (2016) 166:755–65. doi: 10.1016/j.cell.2016.05.069
18. Chong J, Yamamoto M, Xia J. MetaboAnalystR 2.0: from raw spectra to biological insights. *Metabolites.* (2019) 9:57. doi: 10.3390/metabo9030057
19. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* (2010) 26:966–8. doi: 10.1093/bioinformatics/btq054
20. Wei Huang D, Sherman BT, Lempicki RA. *Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources E-mail Addresses for Authors*. (2019). Available online at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.452.9594&rep=rep1&type=pdf (accessed May 3, 2019).
21. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* (2009) 37:1–13. doi: 10.1093/nar/gkn923
22. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* (2018) 46:D649–55. doi: 10.1093/nar/gkx1132
23. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* (2017) 45:D353–61. doi: 10.1093/nar/gkw1092
24. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* (2019) 47:W234–41. doi: 10.1093/nar/gkz240
25. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One.* (2011) 6:e21800. doi: 10.1371/journal.pone.0021800
26. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the pathway interaction database. *Nucleic Acids Res.* (2009) 37:D674–9. doi: 10.1093/nar/gkn653
27. Tong DL, Boocock DJR, Dhondalay GK, Lemetre C, Ball GR. Artificial neural network inference (ANNI): a study on gene-gene interaction for biomarkers in childhood sarcomas. *PLoS One.* (2014) 9:e102483. doi: 10.1371/journal.pone.0102483
28. Abdel-Fatah TMA, Agarwal D, Liu D-X, Russell R, Rueda OM, Liu K, et al. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. *Lancet Oncol.* (2016) 17:1004–18. doi: 10.1016/S1470-2045(16)00174-1
29. Zafeiris D, Rutella S, Ball GR. An artificial neural network integrated pipeline for biomarker discovery using alzheimer's disease as a case study. *Comput Struct Biotechnol J.* (2018) 16:77–87. doi: 10.1016/j.csbj.2018.02.001
30. Nassiri F, Mamatjan Y, Suppiah S, Badhiwala JH, Mansouri S, Karimi S, et al. DNA methylation profiling to predict recurrence risk in meningioma: development and validation of a nomogram to optimize clinical management. *Neuro Oncol.* (2019) 21:901–10. doi: 10.1093/neuonc/noz061
31. Bi WL, Abedalthagafi M, Horowitz P, Agarwalla PK, Mei Y, Aizer AA, et al. Genomic landscape of intracranial meningiomas. *J Neurosurg.* (2016) 125:525–35. doi: 10.3171/2015.6.JNS15591
32. Fan C, Yu J, Liu Y, Xu H, Wang E. Increased NDRG1 expression is associated with advanced T stages and poor vascularization in non-small cell lung cancer. *Pathol Oncol Res.* (2012) 18:549–56. doi: 10.1007/s12253-010-9294-2
33. Azuma K, Kawahara A, Hattori S, Taira T, Tsurutani J, Watari K, et al. NDRG1/Cap43/Drg-1 may predict tumor angiogenesis and poor outcome in patients with lung cancer. *J Thorac Oncol.* (2012) 7:779–89. doi: 10.1097/JTO.0b013e31824c92b4
34. Wilisch-Neumann A, Kliese N, Pachow D, Schneider T, Warnke J-P, Braunsdorf WE, et al. The integrin inhibitor cilengitide affects meningioma cell motility and invasion. *Clin Cancer Res.* (2013) 19:5402–12. doi: 10.1158/1078-0432.CCR-12-0299
35. Wang X, Gong Y, Wang D, Xie Q, Zheng M, Zhou Y, et al. Analysis of gene expression profiling in meningioma: deregulated signaling

pathways associated with meningioma and EGFL6 overexpression in benign meningioma tissue and serum. *PLoS One.* (2012) 7:e52707. doi: 10.1371/journal.pone.0052707
36. Campbell ID, Humphries MJ. Integrin structure, activation, and interactions. *Cold Spring Harb Perspect Biol.* (2011) 3:a004994. doi: 10.1101/cshperspect.a004994
37. Hannigan G, Troussard AA, Dedhar S. Integrin-linked kinase: a cancer therapeutic target unique among its ILK. *Nat Rev Cancer.* (2005) 5:51–63. doi: 10.1038/nrc1524
38. Nigim F, Wakimoto H, Kasper E, Ackermans L, Temel Y. Emerging medical treatments for meningioma in the molecular era. *Biomedicines.* (2018) 6:86. doi: 10.3390/biomedicines6030086
39. Okamoto H, Li J, Vortmeyer AO, Jaffe H, Lee Y-S, Gläsker S, et al. Comparative proteomic profiles of meningioma subtypes. *Cancer Res.* (2006) 66:10199–204. doi: 10.1158/0008-5472.CAN-06-0955
40. Bouamrani A, Ramus C, Gay E, Pelletier L, Cubizolles M, Brugiè Re S, et al. Increased phosphorylation of vimentin in noninfiltrative meningiomas. *PLoS One.* (2010) 5:e9238. doi: 10.1371/journal.pone.0009238
41. Satelli A, Li S. Vimentin in cancer and its potential as a molecular target for cancer therapy. *Cell Mol Life Sci.* (2011) 68:3033–46. doi: 10.1007/s00018-011-0735-1
42. Reddy TRK, Li C, Guo X, Fischer PM, Dekker LV. Design, synthesis and SAR exploration of tri-substituted 1,2,4-triazoles as inhibitors of the annexin A2-S100A10 protein interaction. *Bioorganic Med Chem.* (2014) 22:5378–91. doi: 10.1016/j.bmc.2014.07.043
43. Abbritti RV, Polito F, Cucinotta M, Lo Giudice C, Caffo M, Tomasello C, et al. Meningiomas and proteomics: focus on new potential biomarkers and molecular pathways. *Cancer Genomics Proteomics.* (2016) 13:369–79.
44. Cao Y, Wei M, Li B, Liu Y, Lu Y, Tang Z, et al. Functional role of eukaryotic translation initiation factor 4 gamma 1 (EIF4G1) in NSCLC. *Oncotarget.* (2016) 7:24242–51. doi: 10.18632/oncotarget.8168
45. Myrvang HK, Guo X, Li C, Dekker LV. Protein interactions between surface annexin A2 and S100A10 mediate adhesion of breast cancer cells to microvascular endothelial cells. *FEBS Lett.* (2013) 587:3210–5. doi: 10.1016/j.febslet.2013.08.012
46. Grindheim AK, Saraste J, Vedeler A. Protein phosphorylation and its role in the regulation of Annexin A2 function. *Biochim Biophys Acta Gen Subj.* (2017) 1861:2515–29. doi: 10.1016/j.bbagen.2017.08.024
47. Bhavanasi D, Speer KF, Klein PS. CKAP4 is identified as a receptor for Dickkopf in cancer cells. *J Clin Invest.* (2016) 126:2419–21. doi: 10.1172/JCI88620
48. Oblinger JL, Burns SS, Huang J, Pan L, Ren Y, Shen R, et al. Overexpression of eIF4F components in meningiomas and suppression of meningioma cell growth by inhibiting translation initiation. *Exp Neurol.* (2018) 299:299–307. doi: 10.1016/j.expneurol.2017.06.015
49. Pengru H, Yujie C, Bin Z, Lili C. Roles of NUCKS1 in diseases: susceptibility, potential biomarker, and regulatory mechanisms. *Biomed Res Int.* (2018) 2018:7969068. doi: 10.1155/2018/7969068
50. Zhu K, Zhang Z, Zhang H, Wang Z, Wang F. MiR-142-3p targeting NUCKS1 inhibits proliferation and invasion of pancreatic cancer cells. *Artif Cells Nanomed Biotechnol.* (2020) 48:415–24. doi: 10.1080/21691401.2019.1652629
51. Gelman IH. Emerging roles for SSeCKS/Gravin/AKAP12 in the control of cell proliferation, cancer malignancy, and barriergenesis. *Genes Cancer.* (2010) 1:1147–56. doi: 10.1177/1947601910392984
52. Tröger J, Moutty MC, Skroblin P, Klussmann E. A-kinase anchoring proteins as potential drug targets. *Br J Pharmacol.* (2012) 166:420–33. doi: 10.1111/j.1476-5381.2011.01796.x
53. Liu H, Shim AHR, He X. Structural characterization of the ectodomain of a disintegrin and metalloproteinase-22 (ADAM22), a neural adhesion receptor instead of metalloproteinase. Insights on adam function. *J Biol Chem.* (2009) 284:29077–85. doi: 10.1074/jbc.M109.014258
54. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* (2006) 34:D668–72. doi: 10.1093/nar/gkj067
55. GeneCards. *ITGA5 Gene–GeneCards | ITA5 Protein | ITA5 Antibody*. (2020). Available online at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=ITGA5 (accessed June 6, 2020).
56. GeneCards. *AKT1S1 Gene–GeneCards | AKTS1 Protein | AKTS1 Antibody*. (2020). Available online at: https://www.genecards.org/cgi-bin/carddisp.pl?gene=AKT1S1#drugs_compounds (accessed June 6, 2020).
57. Skiriute D, Tamasauskas S, Asmoniene V, Saferis V, Skauminas K, Deltuva V, et al. Tumor grade-related NDRG2 gene expression in primary and recurrent intracranial meningiomas. *J Neurooncol.* (2011) 102:89–94. doi: 10.1007/s11060-010-0291-9

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 © 2020 Mukherjee, Biswas, Gadre, Jain, Syed, Stylianou, Zeng, Mahadevan, Epari, Shetty, Moiyadi, Roy Ball and Srivastava. 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 Oncology

ORIGINAL RESEARCH
published: 28 August 2020
doi: 10.3389/fonc.2020.01698

Toxicoproteomics Disclose Pesticides as Downregulators of TNF-α, IL-1β and Estrogen Receptor Pathways in Breast Cancer Women Chronically Exposed

*Luciana Pizzatti[1], Aedra Carla Bufalo Kawassaki[2], Bruna Fadel[1], Fabio C. S. Nogueira[1], Joseph A. M. Evaristo[1], Nicole Woldmar[1], Géssica Tuani Teixeira[2], Janaína Carla Da Silva[2], Thalita Basso Scandolara[1,2], Daniel Rech[2], Luciano Pessôa Zanetti Candiotto[2], Guilherme Ferreira Silveira[3], Wander Rogério Pavanelli[4] and Carolina Panis[2,4]**

[1] Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, [2] State University of West Paraná, UNIOESTE, Paraná, Brazil, [3] Carlos Chagas Institute, Curitiba, Brazil, [4] State University of Londrina, Londrina, Brazil

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Ali M. Ardekani,
Avicenna Research Institute (ARI), Iran
Anup Kumar Singh,
Beckman Research Institute at City of Hope, United States

***Correspondence:**
Carolina Panis
carolpanis@hotmail.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 13 April 2020
Accepted: 30 July 2020
Published: 28 August 2020

Citation:
Pizzatti L, Kawassaki ACB, Fadel B, Nogueira FCS, Evaristo JAM, Woldmar N, Teixeira GT, Da Silva JC, Scandolara TB, Rech D, Candiotto LPZ, Silveira GF, Pavanelli WR and Panis C (2020) Toxicoproteomics Disclose Pesticides as Downregulators of TNF-α, IL-1β and Estrogen Receptor Pathways in Breast Cancer Women Chronically Exposed. Front. Oncol. 10:1698. doi: 10.3389/fonc.2020.01698

Deleterious effects have been widely associated with chronic pesticide exposure, including cancer development. In spite of several known consequences that pesticides can trigger in the human body, few is known regarding its impact on breast cancer women that are chronically exposed to such substances during agricultural work lifelong. In this context, the present study performed a high-throughput toxicoproteomic study in association with a bioinformatics-based design to explore new putative processes and pathways deregulated by chronic pesticide exposure in breast cancer patients. To reach this goal, we analyzed comparatively non-depleted plasma samples from exposed (n = 130) and non-occupationally exposed (n = 112) women diagnosed with breast cancer by using a label-free proteomic tool. The list of proteins differentially expressed was explored by bioinformatics and the main pathways and processes further investigated. The toxicoproteomic study revealed that women exposed to pesticides exhibited mainly downregulated events, linked to immune response, coagulation and estrogen-mediated events in relation to the unexposed ones. Further investigation shown that the identified deregulated processes and pathways correlated with significant distinct levels tumor necrosis factor alpha and interleukin 1 beta in the blood, and specific clinicopathological characteristics pointed out by bioinformatics analysis as adipose-trophic levels, menopause and intratumoral clots formation. Altogether, these findings reinforce pesticides as downregulators of several biological process and highlight that these compounds can be linked to poor prognosis in breast cancer.

Keywords: breast cancer, toxicoproteomics, tumor necrosis factor alfa, interleukin beta 1, estrogen signaling

INTRODUCTION

Pesticides are chemical compounds widely used in agriculture to control pests since 1960s (1). Regardless of its specific targets, such substances unfortunately reach the human organism, and negative cumulative effects have been reported in people worldwide (2). In this context, chronic exposure to pesticides has been discussed as a significant risk factor for the development of cancer, including breast tumors (3).

Breast cancer is the most frequently malignant neoplasia diagnosed in women worldwide, whose origin is mostly connected to life habits and the environment, and to a lesser extent to inheritable genetic mechanisms (4). Therefore, the contribution of substances that are continuously present as contaminants in the environment, may have a pivotal role in the genesis of breast cancer (3), especially in geographic areas in which women are important players in the rural work and are in constant contact with these compounds (5, 6).

In vitro and experimental studies have reported the mechanisms triggered by pesticides that contribute to breast carcinogenesis, which fall essentially within DNA damage-based events in association with hormones deregulation and rising of metabolites that activate oncogenes (7, 8). However, few is known about how these mechanisms interconnects, as well as its correlation with the disease prognosis and clinicopathological features in human breast cancer as a result of the toxic consequences of pesticide exposure.

In the last years, aiming to expand the knowledge beyond isolated biological findings, high throughput molecular approaches combined with bioinformatics designs has raised as powerful tools to understand tumor behavior and biology. Thereby, it became possible to demonstrate that breast cancer is a challenging disease and have distinct mechanisms activated depending on specific clinicopathological characteristics (9–12).

As far as we know, until now there is no studies reporting the use of proteomics-based strategies to assess the clinical impact of chronic pesticide exposure in women with breast cancer. Relevant information, as how the deregulated biological processes induced by pesticides are interconnected to clinical parameters, are still missing. To fill this gap, the present study proposes a toxicoproteomic perspective to investigate the systemic profile of differentially expressed proteins in blood samples of breast cancer patients chronically exposed to pesticides. By using a high-throughput label-free proteomic strategy, we provide an integrative clinicopathological view based on deregulated events pointed out by bioinformatics approaches and corroborated by further analytical investigations.

MATERIALS AND METHODS

Design of the Study, Patient Selection and Sample Collection

This is a prospective study conducted between May 2015 and December 2018 approved by the Institutional Ethics Board under the number CAAE 35524814.4.0000.0107. The sample size included 242 patients diagnosed with breast cancer attended by the 8th Health Care Region of the State of Paraná at Francisco Beltrão Cancer Hospital, Paraná, Brazil, which corresponds to a total of 27 municipalities. All patients signed consent forms and each protocol followed the principles for medical research involving human subjects from the Declaration of Helsinki. The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) criteria were followed regarding patient selection, assay performance, and data analysis throughout the study.

To characterize the exposure, patients were invited to answer a questionnaire with 61 questions about their current and past occupational history. Based on their answers, we categorized the study population as occupationally exposed or not to pesticides. The criteria of inclusion applied in this study for the exposed group was chronic direct contact with pesticides (pesticide dilution and spraying without personal protection equipment – PPE and/or washing of contaminated clothes and PPE without gloves, at least once a month for the last 40 years). The main objective was to identify whether and how the patient had contact with pesticides: applying it directly or indirectly (by washing clothes or touching equipment and pesticide packs or merely living in areas where pesticides has been used). Unexposed group was composed mainly by women that lived most of their life in urban areas, without any occupational contact with pesticides. Therefore, patients were categorized according to their occupational status as exposed ($n = 130$) and non-occupationally exposed or unexposed ($n = 112$) to pesticides. Patients from both groups were matched for age.

Heparinized blood samples were collected by peripheral venous puncture (5 mL), centrifuged at 4000 rpm for 5 min and plasma aliquots kept frozen at −20°C until analysis. Clinical records were assessed to obtain clinicopathological information. Aiming to avoid any potential bias induced by chemotherapy, all patients included in the study were not under any treatment.

Proteomic Analysis

In-Solution Tryptic Digestion

Proteomics analysis was performed on both groups using the pooled plasma samples of exposed and unexposed patients as strategies previously described (11, 12). Protein concentrations of cleared supernatants were determined using the Qubit™ Protein Assay Kit (Life Technologies). Samples were concentrated and ex-changed with 50 mM ammonium bicarbonate using a 3-kD ultra-filtration device (Millipore). Proteins extracts (50 μg) were then denatured (0.1% RapiGEST at 60°C for 15 min, Waters, Milford, United States), reduced (10 mM dithiothreitol at 60°C for 30 min), alkylated (10 mM iodoacetamide for 30 min at room temperature in the dark), and enzymatically digested with trypsin at a 1:50 (w/w) enzyme to protein ratio. Digestion was terminated by the addition of 10 μL of 5% trifluoroacetic acid (TFA) (13). Peptides were desalted in C18 micro columns (Harvard apparatus), dried in a vacuum centrifuge, resuspended in 0.1% formic acid, quantified by Qubit protein assay, and analyzed by label-free analysis.

LC-MS/MS Analyses

Two μg of digested peptides was analyzed in technical triplicate after 3 h of gradient (5% to 40% B/167 min; 40% to 95% B/5 min; and 95% B/8 min). Easy-nanoLC1000 (Thermo fisher) solvent

A consisted of [95% H2O/5% acetonitrile (ACN)/0.1% formic acid] and solvent B of (95% ACN/5% H2O/0.1% formic acid). Trap-column used was Easy column C18, 2 cm × 100 μm i. d. × 5 μm, 120Å and analytical column of 25 cm and internal diameter of 75 μm (3 μm spheres, Reprosil Pur C18). Label-free quantification was performed in an Easy-nLC 1000 (Thermo Scientific) coupled to a QExactive Plus in FullScan-DDA MS2 mode used a dynamic exclusion list of 45 s and spray voltage at 2.70 kV. Full scan was acquired at a resolution of 70000 at m/z 200, with a m/z range of 350-2000, AGC of 1 × 106, and injection time of 50 ms. Selection of the 15 most intense ions for HCD fragmentation used a normalized collision energy of 30, precursor isolation window of m/z 1.2 and 0.5 offset, a resolution of 17 500 at m/z 200, AGC at 5×10^5, and injection time of 100 ms (14).

Data Analysis

All samples replicate data were analyzed by the Proteome Discoverer 2.1 software using human database UniProt (V. Nov 2018-https://www.uniprot.org/). The parameters used were: full-tryptic search space, up to two missed cleavages allowed for trypsin, precursor mass tolerance of 10 ppm, and fragment mass tolerance of 0.05 Da. Carbamidomethylation of cysteine was included as fixed modification, and methionine oxidation and protein N-terminal acetylation were included as dynamic modifications in label-free quantification. Spectra analyses used a target-decoy strategy considering maximum delta CN of 0.05, all available peptide-spectrum matches, and a target false discovery rate (FDR) 0.01 (strict) as described previously (14). Parameters in the peptide filter were set up for high confidence with a minimum peptide length of six amino acids. For protein filter, we considered the minimum number of peptide sequence as 1, counting only rank 1 peptide. Peptides shared between multiple proteins was counted for the top scoring protein. The confidence thresholds in FDR protein validator were 0.01 for target FDR (strict). The strategy for protein grouping was strict parsimony. Statistical analysis was performed using Perseus Computational Platform v 1.6.10.50. The strategy with both datasets was Log2 transformation followed by subtract median normalization Student *t* test (p-value < 0.05). Volcano plot distribution presented demonstrated the log p-value vs log2 fold differences values. The process and biological pathway validation strategies was previously used and discussed by Pizzatti et al. (13) and Gjertsen and Wiig (15).

In silico Functional Analysis

Functional *in silico* analysis were carried out with FunRich: Functional Enrichment Analysis Tool and MetaCore software using the integrated databases: Human Protein Reference Database (HPRD); Entrez Gene and Uniprot for biological pathways and biological process. For protein-protein interactions the following data bases were accessed: BioGRID; Intact e Human Proteinpedia. For localization, expression data and signaling pathways information the databases: Human Protein Altas; Human Proteome Browser; Human Proteome Map, Proteomics DB, Reactome; NCI; Cell map; and HumanCyc were used. The transcription factor consensus sequences prediction data were obtained with automated search in 29 mammals data bases on the fly in the FunRich tool followed by Bonferroni statistical analysis. Interaction network analysis was evaluated with String functional protein association network database[1] using network analysis based on evidence with minimum required interaction score of 0.400. Venn Diagram the tool[2] was used.

Clinicopathological-Applied Investigation

Based on the results obtained from high-throughput proteomic screening and *in silico* data analysis, we investigated the main downregulated events triggered by chronic pesticide exposure in breast cancer patients individually.

Considering that most of the processes and pathways were downregulated by pesticide exposure, and that several of these events were connected with inflammation and immune response, we chose to measure two central cytokines commonly produced by patients in breast cancer (16, 17), the tumor necrosis factor alpha (TNF-α), and interleukin 1 beta (IL-1β). Analyses were performed by using commercial antibody-specific enzyme-linked immunosorbent assay (ELISA) kits (eBioscience Inc, United States) with internal controls. Results were calculated in pg/mL by fitting the data to a standard curve obtained using human recombinant cytokines.

Aiming to perform further investigation based on processes and pathways pointed out by bioinformatics, TNF-α and IL-1β results were categorized according to clinicopathological parameters of patients, which included age at diagnosis (cut-off at 50 years), tumor histological grade (I, II, and III), estrogen (ER), and progesterone (PR) receptors expression, lymph node metastasis presence or absence, intratumoral clots presence or absence, menopausal status at diagnosis, tumor size ($\leq$2 cm, $>$2 cm, and $<$5 cm, $\geq$5 cm), ki-67 proliferation index (cut-off at 14%), and trophic-adipose levels based on body mass index (eutrophic $\leq$24.9 kg/m2, overweight between 25.0 and 29.9 kg/m2, and obese $\geq$ 30.0 kg/m2).

Statistical Analysis

Experiments were carried out in duplicate datasets. Results were analyzed by using the Grubbs test for outlier detection and compared by Student's *t*-test, Mann–Whitney or ANOVA, according to variances distribution and the number of the groups compared. P-value < 0.05 was considered significant. All statistical analyses were performed using GraphPad Prism version 7.0 (GraphPad Software, San Diego, CA, United States).

RESULTS

Detailed clinicopathological data of all patients included in the study are shown in **Supplementary Table 1**. Both groups presented similar ages at diagnosis and homogeneous distribution regarding other clinicopathological parameters. The mean age at diagnosis for unexposed patients was 55.2 years, ranging from 33 to 81 years, while the mean age at diagnosis

[1]https://string-db.org/

[2]http://bioinformatics.psb.ugent.be/webtools/Venn/

for the occupationally exposed patients was 54.6 years, ranging from 30 to 86 years (please see the figure below). The comparison between groups retrieved a not significant *p* value of 0.7570 (Student's *t*-test). No statistical differences were found when comparing each specific parameter between both groups.

Regarding the characterization of pesticide exposure, exposed women reported that they lived at least 50% of their lives working with pesticides, and to work at least once a week in direct contact with pesticides, as the following: 1. washing the clothes and personal protection equipment impregnated with pesticides worn by family members who applied these substances, 2. Preparing and diluting the concentrated pesticides, 3. Helping the spraying of diluted pesticides in the crops (time estimated: 4–8 h per day, during 2–3 consecutive days, each 1–2 weeks). Considering that Brazil has practically 100% of its arable land and that the area of study is one of the major regions of agriculture in Paraná State (second gross domestic product of the State), the occupational exposure of such patients is very intense. Further, 94% of interviewed women from exposed group reported to perform all activities without wearing personal protection equipment, not even gloves. Taking into consideration that pesticides are majorly absorbed by skin, and the duration/chronicity of the contact of patients with these substances, this is a worthwhile route of contamination, bigger than any other source as food or water. On the other hand, not occupationally exposed women did not exert any type of rural work enrolling pesticides, neither have any occupational contact with pesticides. When they answered the questionnaire, they reported none occupational contact (current or past) with pesticides. These women do not have the history of washing clothes or personal protection equipment contaminated with pesticides, and never had to apply pesticides in the crops. Therefore, they were selected to the unexposed group due to their completely distinct profile of occupational exposure history.

It is important to highlight that food consumption habits are similar in both exposed and unexposed groups, and the water system that supplies the study area is the same. Whether it is contaminated with pesticides, the occupational exposure overlaps it, due to its intensity and severity in the exposed group. Considering the intensity of chronic occupational exposure to pesticides, we considered that there is substantial information about pesticide exposure in the population of the study, and that the exposed patients are under completely distinct conditions if compared to the unexposed ones, independent on their contact with food and water pesticides (that are the same levels for both groups).

The proteomic screening identified in exposed and unexposed replicates a total of 554 proteins, with 280 proteins in exposed and 274 in unexposed ones. Information about total protein ratios and its raw data are reported in **Supplementary Tables 2, 3**.

The Venn diagram (**Figure 1A**) shows the qualitative distribution of the identified proteins: 200 proteins identified in both conditions: 32 proteins identified in exclusively in exposed samples: and 8 proteins identified exclusively in unexposed samples. Regarding the differentially expressed proteins when comparing the groups, a total of 142 proteins were found upregulated and 41 downregulated in the exposed ones (**Figure 1B**).

The histogram distribution obtained after a Log2 transformation and subtract median normalization is shown in **Figure 1C**. General Label-free quantification analysis between the proteomic profiles identified 183 proteins differentially expressed. 142 proteins with exposed/unexposed ratio ≥ 1 and 41 proteins with exposed/unexposed ratio ≤ 1 (**Supplementary Table 2** and **Figures 1E,F**). Between the initial 554 identified proteins in both groups, 243 protein presented quantified valid values and were present in three of three sample replicates in both groups. After the statistical analysis in Perseus software and appropriated filtering, 42 proteins were found upregulated and 8 proteins were found downregulated as shown in the volcano plot (**Figure 1D**). In **Table 1**, the list of differentially proteins showed in the volcano plot are presented with the respective Log *P*-values and Log2 fold differences.

The *in silico* analysis (**Figure 2**) performed in FunRich: Functional Enrichment Analysis Tool software revealed the major biological events in which all the up and downregulated proteins differently expressed in the chronically pesticide-exposed patients are enrolled. **Figure 2A** shows an amount of 15 biological pathways in which 11 are found downregulated in the chronically pesticide-exposed patients. These pathways were related to adhesion (5), receptors and second messengers signaling (5), immune response (3), and coagulation mechanisms (2). Most of them are intrinsically connected with immune-mediated events. In respect to the 13 biological processes identified through the bioinformatics evaluation (**Figure 2B**), 5 were predominantly downregulated in the chronically pesticide-exposed group of patients. The processes affected were related to metabolism and energy, protein metabolism, cell growth, and apoptosis.

Using the Funrich data base for transcription factors consensus binding sites sequences, the prediction of potential transcriptional factors related (direct or indirectly) with the identified protein data sets, were predicted. In **Figure 2C** is showed the top nine transcription factor (DNA consensus sequences) that are present in the promoter sequences of the identified proteins. It is noted that most of the proteins were related to Hepatocyte nuclear factor 1-alpha HNF1A (in downregulated dataset), HNF4A Hepatocyte nuclear factor 4-alpha (in upregulated data set), Estrogen-related receptor alpha (ESRRA; in downregulated data set) Forkhead box protein A1 FOXA1 (in downregulated data set) and Signal transducer and activator of transcription 1-alpha STAT1 (upregulated data set).

Moreover, the functional protein association networks were analyzed. In String data base software, using network analysis based on evidence with minimum required interaction score of 0.400 using its database. The list of proteins used to perform this analysis is shown in **Supplementary Table 2** and in **Table 1**. The results of the network analysis retrieved from String database software is shown in **Supplementary Tables 4, 5**. The functional enrichments interactions observed after the *in silico* analysis of the up and down regulated datasets showed an enrichment *p*-value of $< 1.0e\text{-}16$ and are presented in the **Figure 3**. This means that proteins presented in the identified data set have more interactions among themselves than what would be expected for a random set of proteins of similar size, drawn from the

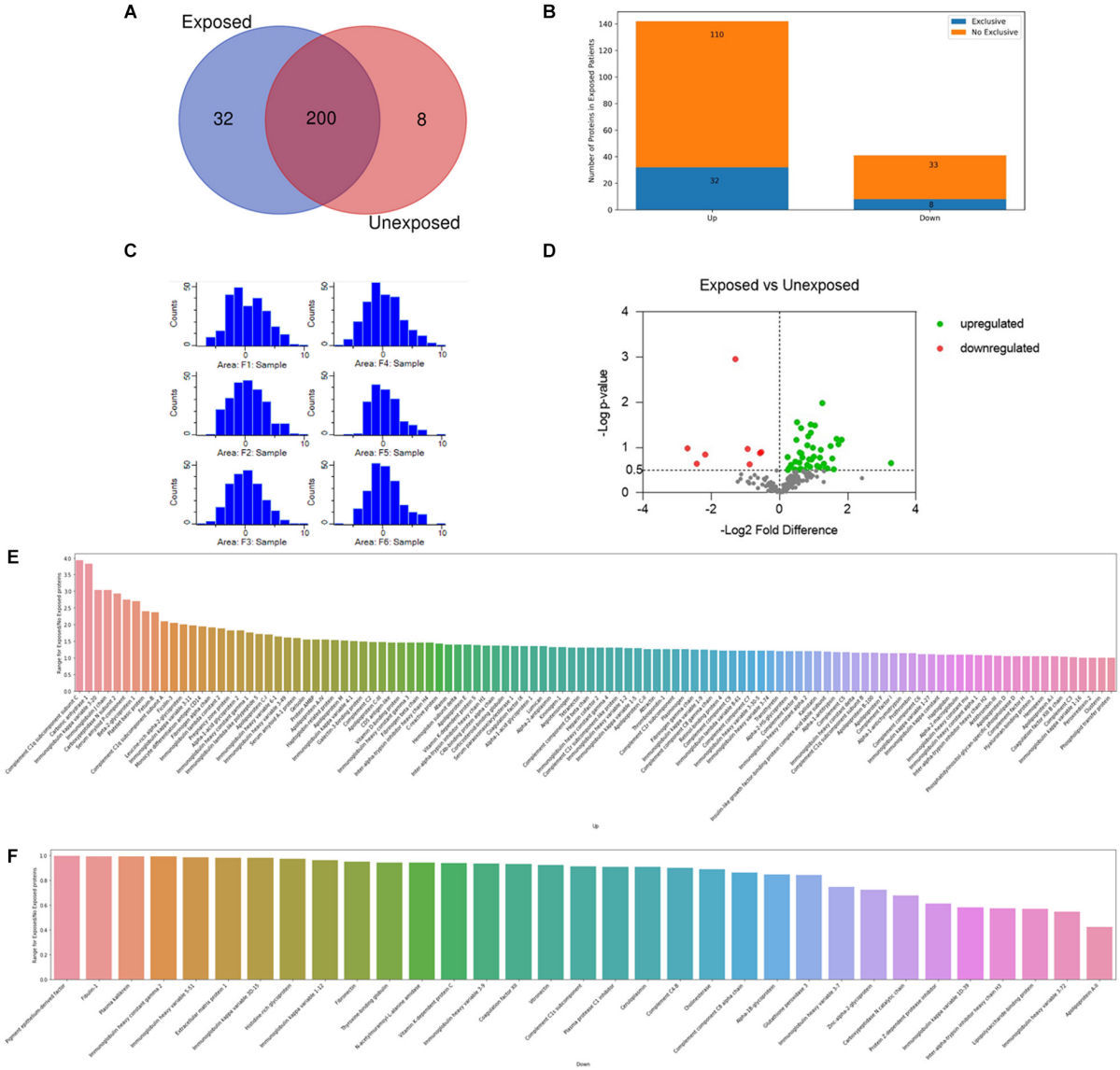


FIGURE 1 | Proteouiic screening results. **(A)** Venn diagram showing the qualitative distribution of the identified proteins in both groups, 200 proteins identified in both conditions, 32 proteins identified in exclusively in exposed samples, and 8 proteins identified exclusively in unexposed samples. **(B)** The bars indicate the number of up and downregulated proteins differentially expressed in the exposed patients. Regarding the differentially expressed proteins when comparing the groups: a total of 142 proteins were found upregulated and 41 downregulated in the exposed ones. **(C)** Histogram of distribution after normalization of the replicates. Exposed (Fl, F2, and F3) and unexposed (F4, F5, and F6). **(D)** Volcano plot of the differentially expressed proteins between exposed and unexposed samples. *In* green and red upregulated and downregulates proteins respectivelly. Tn gray proteins with -log *p*-value < 0.5. Total proteins identified as up **(E)** or downregulated **(F)** are reported accordingly to their specific ratios.

genome. Such an enrichment indicates that the proteins are at least partially biologically connected, as a group. For the downregulated protein dataset (**Figure 3A**), this *in silico* analysis also showed as top 5 Biological process (GO) hits: regulation of inflammatory response (false discovery rate of 1.82e-11); protein activation cascade (false discovery rate 9.32e-11); regulation of peptidase activity (false discovery rate 1.11e-10); and regulation of response to external stimulus (false discovery rate 1.15e-10) as major biological process interaction. The major biological process identified in the network interaction of the upregulated identified dataset (**Figure 3B**) are regulation of complement activation (false discovery rate of 2.59e-35); regulation of protein processing (false discovery rate 1.19e-34) and regulation of humoral immune response (false discovery rate 1.19e-34).

According to the results from the *in silico* study we performed a search in the literature to understand the putative connections among the highlighted processes and pathways downregulated in breast cancer patients by chronic pesticide exposure. Based on this, it was possible to note that most of data were linked to inflammation and immune-related events. Therefore, aiming

TABLE 1 | Differentially expressed proteins in chronically pesticide-exposed patients. Volcano plot data.

Accession (UniProt)	Description	−Log2 fold difference	−Log *p*-value
Down regulated proteins in exposed samples			
Q15848	Adiponectin OS = Homo sapiens OX = 9606 GN = ADIPOQ PE = 1 SV = 1	−0,577652	0,877936
P02652	Apolipoprotein A-II OS = Homo sapiens OX = 9606 GN = APOA2 PE = 1 SV = 1	−1,29743	2,95751
P15169	Carboxypeptidase N catalytic chain OS = Homo sapiens OX = 9606 GN = CPN1 PE = 1 SV = 1	−0,5372	0,903306
P12259	Coagulation factor V OS = Homo sapiens OX = 9606 GN = F5 PE = 1 SV = 4	−2,18176	0,848073
A0A0B4J1Y9	Immunoglobulin heavy variable 3-72 OS = Homo sapiens OX = 9606 GN = IGHV3-72 PE = 3 SV = 1	−0,937188	0,972346
A0A0A0MT36	Immunoglobulin kappa variable 6D-21 OS = Homo sapiens OX = 9606 GN = IGKV6D-21 PE = 3 SV = 1	−2,69985	0,985784
P01721	Immunoglobulin lambda variable 6-57 OS = Homo sapiens OX = 9606 GN = IGLV6-57 PE = 1 SV = 2	−2,4305	0,645198
P18428	Lipopolysaccharide-binding protein OS = Homo sapiens OX = 9606 GN = LBP PE = 1 SV = 3	−0,877242	0,632545
Up regulated proteins in exposed samples			
P02763	Alpha-1-acid glycoprotein 1 OS = Homo sapiens OX = 9606 GN = ORM1 PE = 1 SV = 1	0,282142	0,549393
P19652	Alpha-1-acid glycoprotein 2 OS = Homo sapiens OX = 9606 GN = ORM2 PE = 1 SV = 2	0,672468	0,780538
P01019	Angiotensinogen OS = Homo sapiens OX = 9606 GN = AGT PE = 1 SV = 1	0,598796	0,51995
P06727	Apolipoprotein A-IV OS = Homo sapiens OX = 9606 GN = APOA4 PE = 1 SV = 3	0,811802	0,619374
P02655	Apolipoprotein C-II OS = Homo sapiens OX = 9606 GN = APOC2 PE = 1 SV = 1	0,237511	0,792502
P02749	Beta-2-glycoprotein 1 OS = Homo sapiens OX = 9606 GN = APOH PE = 1 SV = 3	1,20994	0,951809
P04003	C4b-binding protein alpha chain OS = Homo sapiens OX = 9606 GN = C4BPA PE = 1 SV = 2	0,446286	0,52209
P00915	Carbonic anhydrase 1 OS = Homo sapiens OX = 9606 GN = CA1 PE = 1 SV = 2	1,67218	1,19446
P00918	Carbonic anhydrase 2 OS = Homo sapiens OX = 9606 GN = CA2 PE = 1 SV = 2	0,632465	1,43293
P22792	Carboxypeptidase N subunit 2 OS = Homo sapiens OX = 9606 GN = CPN2 PE = 1 SV = 3	1,25728	1,98658
P04040	Catalase OS = Homo sapiens OX = 9606 GN = CAT PE = 1 SV = 3	0,913062	1,51147
O43866	CD5 antigen-like OS = Homo sapiens OX = 9606 GN = CD5L PE = 1 SV = 1	0,583446	0,511082
P02745	Complement C1q subcomponent subunit A OS = Homo sapiens OX = 9606 GN = C1QA PE = 1 SV = 2	0,816096	1,05262
P02747	Complement C1q subcomponent subunit C OS = Homo sapiens OX = 9606 GN = C1QC PE = 1 SV = 3	3,27164	0,659602
P0C0L4	Complement C4-A OS = Homo sapiens OX = 9606 GN = C4A PE = 1 SV = 2	1,03893	1,49116
P07360	Complement component C8 gamma chain OS = Homo sapiens OX = 9606 GN = C8G PE = 1 SV = 3	0,24517	0,503356
Q9UGM5	Fetuin-B OS = Homo sapiens OX = 9606 GN = FETUB PE = 1 SV = 2	1,40429	0,541632
P02671	Fibrinogen alpha chain OS = Homo sapiens OX = 9606 GN = FGA PE = 1 SV = 2	1,10486	0,59767
O75636	Ficolin-3 OS = Homo sapiens OX = 9606 GN = FCN3 PE = 1 SV = 2	1,00807	0,802906
Q08380	Galectin-3-binding protein OS = Homo sapiens OX = 9606 GN = LGALS3BP PE = 1 SV = 1	0,465231	0,6933
P06396	Gelsolin OS = Homo sapiens OX = 9606 GN = GSN PE = 1 SV = 1	0,511214	1,56289
P00739	Haptoglobin-related protein OS = Homo sapiens OX = 9606 GN = HPR PE = 2 SV = 2	1,34752	0,539945
P01857	Immunoglobulin heavy constant gamma 1 OS = Homo sapiens OX = 9606 GN = IGHG1 PE = 1 SV = 1	0,585063	0,669012
P01742	Immunoglobulin heavy variable 1-69 OS = Homo sapiens OX = 9606 GN = IGHV1-69 PE = 1 SV = 2	0,64299	0,872349
A0A0B4J1V2	Immunoglobulin heavy variable 2-26 OS = Homo sapiens OX = 9606 GN = IGHV2-26 PE = 3 SV = 1	0,33751	0,617746
P04433	Immunoglobulin kappa variable 3-11 OS = Homo sapiens OX = 9606 GN = IGKV3-11 PE = 1 SV = 1	1,30189	0,644815
P01619	Immunoglobulin kappa variable 3-20 OS = Homo sapiens OX = 9606 GN = IGKV3-20 PE = 1 SV = 2	1,54984	0,759295
A0A0C4DH55	Immunoglobulin kappa variable 3D-7 OS = Homo sapiens OX = 9606 GN = IGKV3D-7 PE = 3 SV = 5	1,82442	1,17608
P06312	Immunoglobulin kappa variable 4-1 OS = Homo sapiens OX = 9606 GN = IGKV4-1 PE = 1 SV = 1	0,870637	0,741953
P0DOY2	Immunoglobulin lambda constant 2 OS = Homo sapiens OX = 9606 GN = IGLC2 PE = 1 SV = 1	0,73101	0,509415
P01703	Immunoglobulin lambda variable 1-40 OS = Homo sapiens OX = 9606 GN = IGLV1-40 PE = 1 SV = 2	0,980606	0,997284
P01700	Immunoglobulin lambda variable 1-47 OS = Homo sapiens OX = 9606 GN = IGLV1-47 PE = 1 SV = 2	1,59502	0,521646
P01701	Immunoglobulin lambda variable 1-51 OS = Homo sapiens OX = 9606 GN = IGLV1-51 PE = 1 SV = 2	0,495094	1,1716
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS = Homo sapiens OX = 9606 GN = ITIH4 PE = 1 SV = 4	0,622172	0,533977
P02750	Leucine-rich alpha-2-glycoprotein OS = Homo sapiens OX = 9606 GN = LRG1 PE = 1 SV = 2	0,941312	0,739914
P08571	Monocyte differentiation antigen CD14 OS = Homo sapiens OX = 9606 GN = CD14 PE = 1 SV = 2	1,18636	0,777581
P02776	Platelet factor 4 OS = Homo sapiens OX = 9606 GN = PF4 PE = 1 SV = 2	0,896119	0,575848
P20742	Pregnancy zone protein OS = Homo sapiens OX = 9606 GN = PZP PE = 1 SV = 4	0,918109	1,33271
P02760	Protein AMBP OS = Homo sapiens OX = 9606 GN = AMBP PE = 1 SV = 1	1,17521	0,568105
Q92954	Proteoglycan 4 OS = Homo sapiens OX = 9606 GN = PRG4 PE = 1 SV = 3	0,841435	1,25299
P0DJI8	Serum amyloid A-1 protein OS = Homo sapiens OX = 9606 GN = SAA1 PE = 1 SV = 1	0,591495	0,891718
P02743	Serum amyloid P-component OS = Homo sapiens OX = 9606 GN = APCS PE = 1 SV = 2	1,50171	1,03638
P62979	Ubiquitin-40S ribosomal protein S27a OS = Homo sapiens OX = 9606 GN = RPS27A PE = 1 SV = 2	1,73161	1,07562

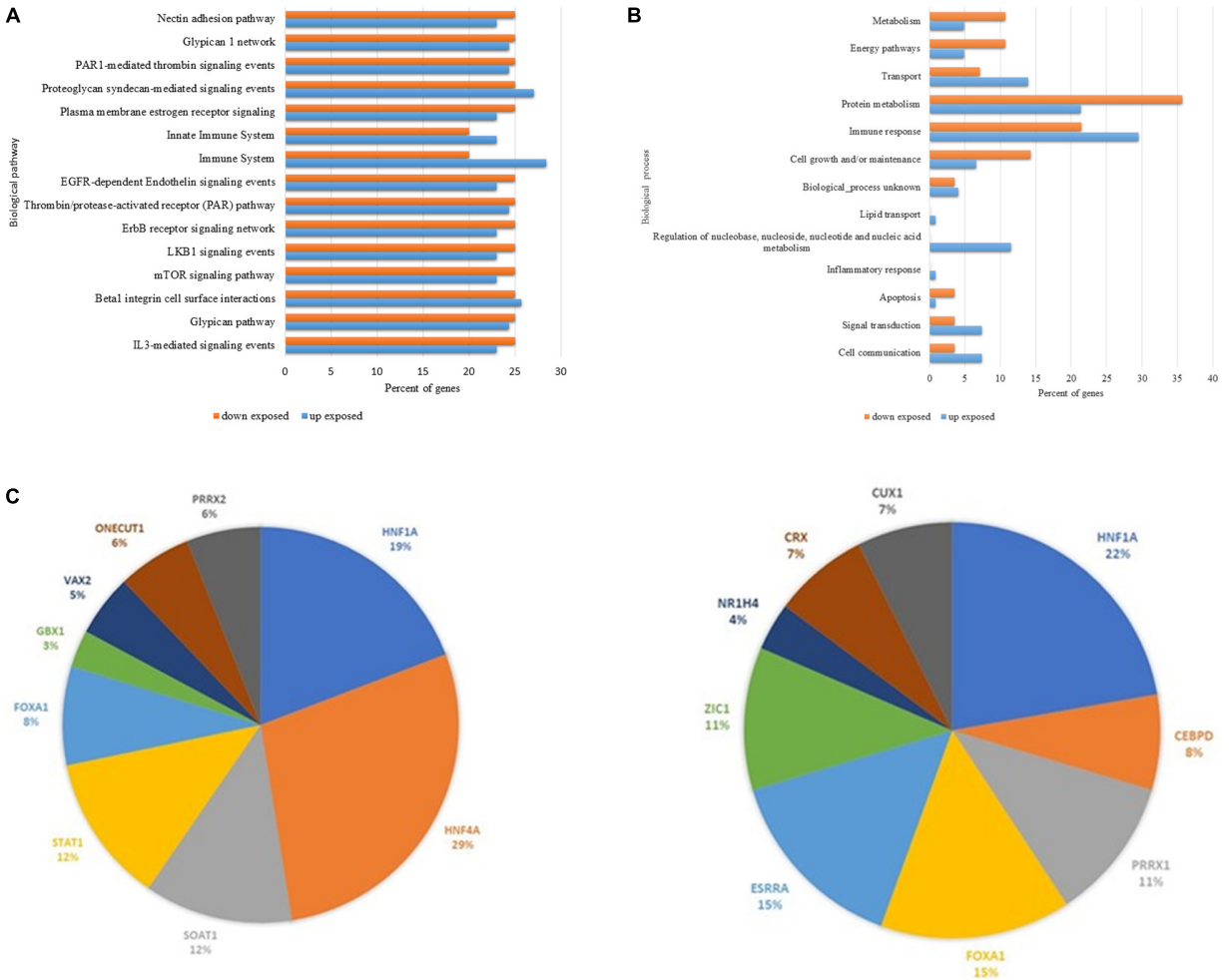


FIGURE 2 | *In stiico* comparative analysis between the main up and down regulated proteins differentially expressed by breast cancer patients chronically exposed to pesticides. **(A)** Biological pathways. **(B)** Biological processes and **(C)** Prediction of potential Transcription fator related to proteins data set regulation. Left panel prediction for up regulated proterins. Right panel prediction for downregulated proteins. The orange columns represent the downregulated events, and the blue columns indicate the upregulated ones. Note that for most of the events, the downregulation is more frequent.

to understand the clinicopathological meaning of our findings and validate the biological relevance and data reliability, we decided to investigate TNF-α and IL-1β levels in plasma samples from both exposed and unexposed groups. TNF-α and IL-1β represents major key factor of inflammation and immune-related events and process and pathway validation strategies, as also previously described (13, 18). These cytokines were chosen due to its well-known production in breast cancer patients (16, 17). To validate the downregulated biological events, these cytokines levels were analyzed considering both general levels and some clinicopathological parameters related to the particular processes and pathways revealed by the *in silico* study. Cytokine levels distributed according to the investigated clinicopathological parameters are reported in **Supplementary Tables 6, 7**.

Tumor necrosis factor alpha levels were significantly reduced in the chronically pesticide-exposed patients (**Figure 4**) when compared to those unexposed (94.31 ± 6.02 pg/mL for exposed and 122.6 ± 10.78 pg/mL for unexposed, $p = 0.0378$). IL-1β levels did not differ (59.24 ± 4 pg/mL for exposed and 71.1 ± 7.62 pg/mL for unexposed, $p = 0.3067$).

To explore the plasma membrane estrogen receptor signaling downregulation, as well the involvement of the downregulated transcription factor ESRRA pointed out by *in silico* analysis, cytokine levels were analyzed according to patients' menopausal status at diagnosis. TNF-α levels (**Figure 5A**) was significantly reduced in breast cancer patients that are chronically exposed to pesticides when compared to the unexposed ones (94.6 ± 8.52 pg/mL and 127.2 ± 13.99 pg/mL, respectively, $p = 0.0443$), while IL-1β was not different.

Concerning the study of metabolism and energy pathways downregulation (**Figure 5B**) suggested by *in silico* analysis, IL-1β levels revealed significantly reduced in the plasma of obese breast cancer patients chronically exposed to pesticides (42.78 ± 4.64 ρg/mL to the exposed and 85.10 ± 14.52 ρg/mL to the unexposed, $p = 0.0247$). For this pathway, TNF-α levels did not change. In relation to the validation of the downregulated

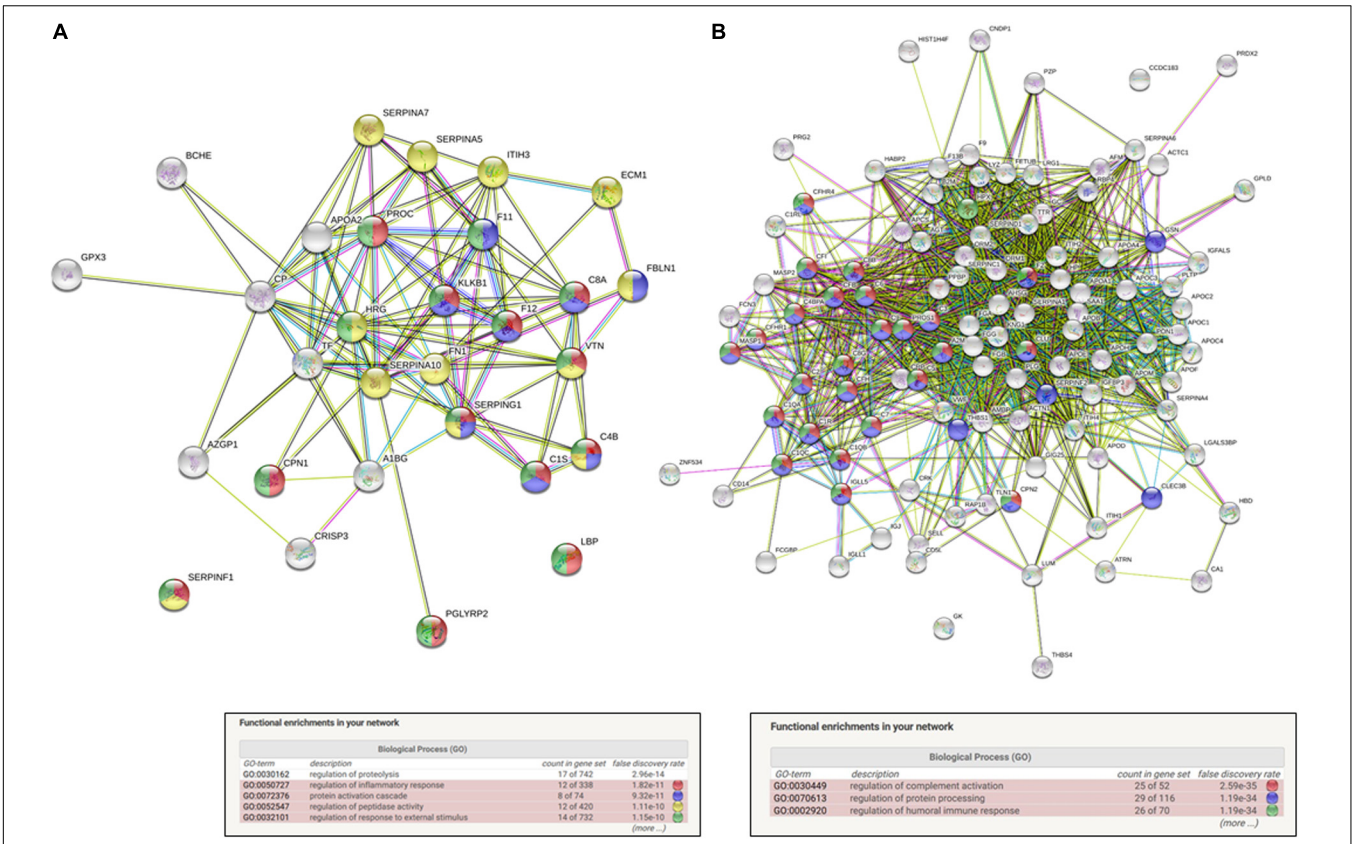


FIGURE 3 | *In silico* functional interaction enrichement network analysis of differentially expressed proteins in breast cancer patients chronically exposed to pesticides. **(A)** Downregulated dataset interaction networks, number of nodes: 25; number of edges: 75; average node degree: 6; avg. local clustering coefficient: 0.605; and expected number of edges: 3: PPI enrichment *p*-value: <1.0e-16. **(B)** Up regulated dataset interaction networks. Number of nodes: 32 number of edges; 109 average node degree; 6.81 avg. local; clustering coefficient: 0.611; expected number of edges: 12; and PPI enrichment p-value: <10e-16.

PAR-1-mediated thrombin signaling, a clot formation-related pathway also highlighted in the *in silico* processes analysis, TNF-α levels were significantly reduced in chronically pesticide-exposed patients that presented intratumoral clots (**Figure 5C**, 89.82 ± 11.31 pg/mL for exposed patients and 151.0 ± 20.99 pg/mL for non-exposed patients, $p = 0.0128$). In this case, IL-1β has had no variation.

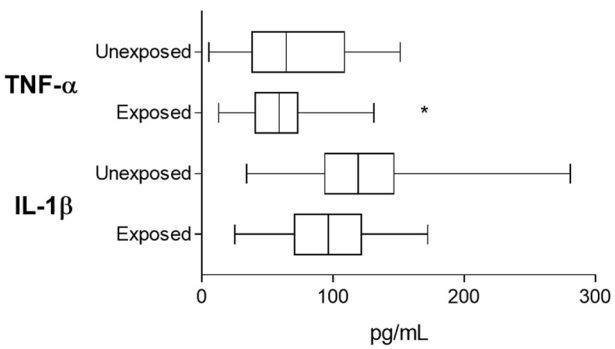


FIGURE 4 | Plasmatic TNF-α and IL-Iβ measurements for validation of imune response related events as pointed out by proteomic screening. Based on the results obtained from proteomic screening that shown the downregulation of imune response and inflammation-related events, the validation step investigated two major cytokines reported as produced in breast cancer patients. *indicates statistical difference ($p < 0.05$). Student's *t* Test.

DISCUSSION

Toxicoproteomics is the field of proteomics that aims to understand the impact of environmental exposures on proteins, processes and pathways in biological systems by using global protein expression approaches (19). Based on this concept, the present study was designed to comprehend the influence of chronic pesticide exposure in systemic proteomic profile of breast cancer patients. Our results demonstrated that these patients have systemically deregulated processes in relation to the unexposed ones, pointing out pesticide exposure as significant downregulators of several biological networks.

Taking into account that such pathways are substantially linked to inflammatory and immune responses, we focused our validation analyses on the clinical impact of pesticide chronic exposure in breast cancer patients regarding their systemic cytokines profile. Furthermore, we pinpointed this scenario based

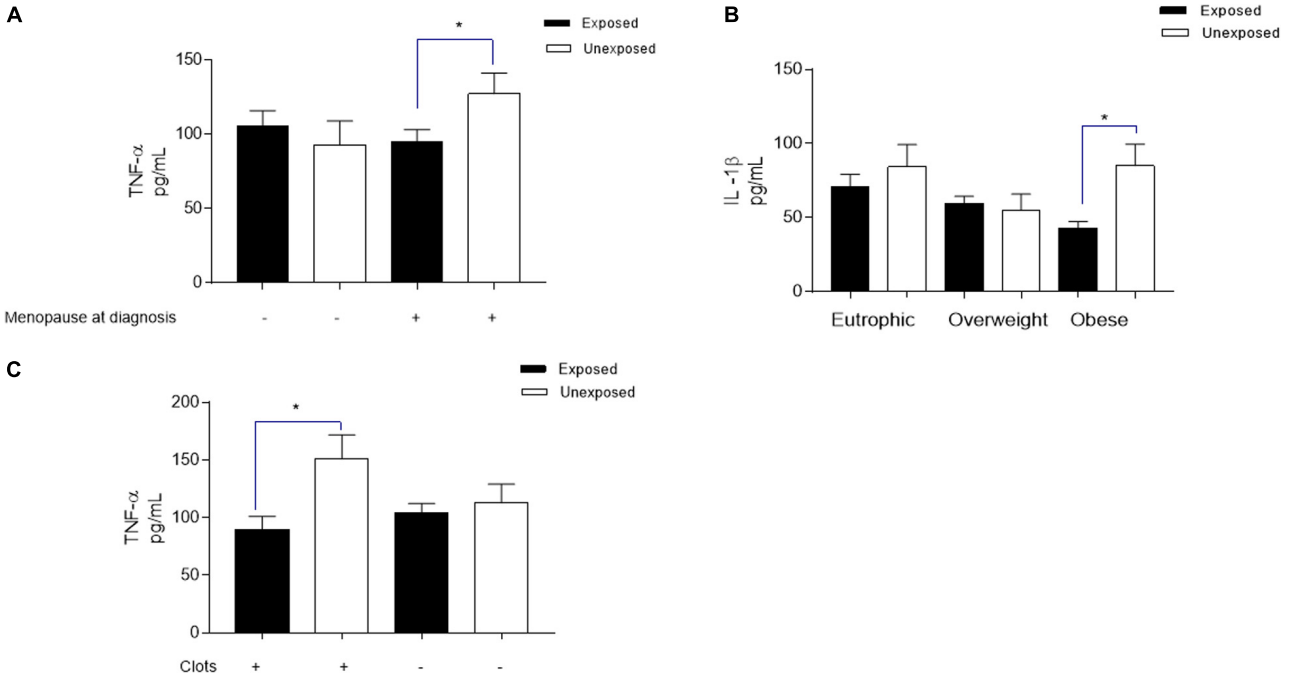


FIGURE 5 | Significant variations in plasmatic TNF-α and IL-Iβ levels from breast cancer patients exposed or unexposed to pesticides according to specific clinicopathological profiles. Clinicopathological parameters investigated were chosen based on the downregulated results obtained from biological pathways, biological processes, and molecular functions revealed by *in silica* analysis in breast cancer cancer chronically exposed to pesticides. For all parameters, the levels of both cytokines were investigated; and here are represented only those with statistical significance for each cytokine. **(A)** For investigation of the plasma membrane estrogen receptor signaling and the transcription factor estrogen-related receptor alpha (ESRRA) pointed out by *in silico* analysis as downregulated (identified in the biological pathways and TFs analysis), the measurement of TNF-ct levels distributed according to patients menopausal status at diagnosis. **(B)** Regarding the metabolism and energy pathways downregulation (identified in the biological processes analysis), the measurement of EL-Iβ levels in plasma of breast cancer patients according to their trophic-adipose status, categorized by ther body mass index as Eutrophic, Overweight and obese. **(C)** To validate the downregulated PAR-1-mediated thrombin signaling events (identified in the biological pathways analysis), TNF-α levels were categorized according to the formation of intratumoral clots observed during biopsies microscopic analysis. *indicates statistical difference ($p < 0.05$).

on a set of prognostic parameters that correlates with the data supported by the toxicoproteomic-based results.

Cytokine production is a central mechanism triggered by immune response against cancer (20), and high levels of systemic TNF-α and IL-1β has been reported in breast cancer patients (16, 17). Our findings demonstrated that chronic exposure to pesticides impairs the systemic cytokine production in breast cancer patients, which leads to lower levels of both TNF-α and IL-1β under specific clinicopathological conditions, when compared to the unexposed group. Cytokine shutdown may result in immunosuppression, a process already documented as a consequence of chronic pesticide exposure (21) that may favor not only cancer development but also enhance its aggressiveness (22). Several pesticides are proven to be immunotoxicants, and a variety of deleterious mechanisms can be listed, including reduction in the number and function of immune cells (23), genetic damage in lymphocytes (24) and suppression of Th1 responses (25). TNF-α can control the activity of HNF1 (26), which helps to understand why it was referred as the main transcription factor downregulated by chronic pesticide exposure in our study.

Nevertheless, pesticides are known for their capability to cause endocrine disruption, which seems to contribute substantially for breast cancer development (27). We found that the most clinically relevant biological pathway pinpointed as downregulated by pesticide exposure in breast cancer patients by *in silico* was the estrogen receptor (ER) signaling, in association with the ESSRA transcription factor. The knockdown of ER axis in breast cancer patients could contribute to the development of the most aggressive phenotype of breast cancer, the triple negative (TNBC). The underlying biological mechanisms that drive TNBC development remains unclear, but ER loss seems to have a role in the metastatic processes. Studies have reported divergences between primary tumors and its metastases, highlighting that the loss of ER is a common event in breast cancer (28, 29), turning favorable prognosis luminal tumors into poor prognosis TNBC. Since pesticide exposure is a common event worldwide, it could be investigated as a possible mechanism in TBNC genesis. It is worth to mention that in this study there are TNBC patients in the exposed group in comparison to the unexposed one, reinforcing these findings.

Since ER signaling was downregulated in patients by pesticide exposure, we have advanced our investigation to determine if cytokine levels were differentially distributed in exposed patients according to their hormonal status. Pesticide exposure did not promote any significant change when comparing the

non-menopausal group of patients. However, exposed patients presented significantly reduced TNF-α levels when compared to the unexposed, suggesting that the exposure may affects TNF-α production in the absence of estrogen. It is known that estrogen modulates Th1/Th17 immune responses (30), and can control the secretion of TNF-α by macrophages (31). Moreover, it is expected that during menopause an increase in TNF-α production occurs as a response to estrogen deprivation, as well as a compensatory mechanism against the decreasing of immune cells (32). However, we did not observe any of these situations in pesticide-exposed patients. Thus, we may conclude that disruptions in TNF-α antitumor mechanisms could result in disease aggravation for patients under pesticide exposure in the future. In this context, we identified that 3.3% of the patients from the unexposed group had recurrence of breast cancer, while about 19% of the occupationally exposed patients recurred. Therefore, a follow-up study is necessary to understand whether recurrence is related to cytokine levels.

Bringing together the results of metabolism and energy processes, both downregulated by *in silico* analysis, IL-1β levels was found significantly reduced in obese patients exposed to pesticides in comparison to the unexposed obese patients. Pesticides affects macrophages function by reducing their lysosomal activity, promoting negative regulation of IL-1β secretion. Additionally, pesticides appear to modulate IL-1β levels in the spleen and thymus of mice fed with high-fat diet (33), and promote its spontaneous secretion by human blood monocytes blood (34), which could lead into the exhaustion of this system. Considering that obesity constitutes a major risk factor for the worst prognosis of breast cancer (35), the reduction of IL-1β in these patients induced by pesticides could constitute an additional aggravating factor.

Among this context of inflammation mechanisms deregulated due to pesticides exposure, we also identified downregulation of components from the coagulation pathway. We observed significantly reduced TNF-α levels in breast cancer patients that exhibited intratumoral clots and were exposed to pesticides. This finding suggests an interplay between TNF-α and HNF1 induced by pesticides in breast cancer patients, since HNF1 directly controls the transcription of blood clotting genes (36) and both were downregulated accordingly to the *in silico* analysis.

In conclusion, the validation experiments performed by crossing cytokines levels and specific clinicopathological parameters, that were chosen based on the bioinformatics analysis, reinforced that toxicoproteomics is a reliable approach to investigate the impact of chronic pesticide exposure in breast cancer patients. These findings also support pesticides as critical downregulators of biological responses and mechanisms that can be implicated in breast cancer worse development and progression.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1] partner repository with the dataset identifier PXD018577.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by State University of West Paraná IRB. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LP, BF, FN, JE, and NW: proteomic and bioinformatic analysis, manuscript discussion, and writing. AK, JD, and TS: data analysis, cytokines measurement, clinicopathological data assessment, manuscript discussion, and writing. DR: patients selection, manuscript discussion, and writing. LC: pesticide exposure categorization, manuscript discussion, and writing. GS: bioinformatics, manuscript discussion, and writing. WP and CP: design of the study, data analysis, study supervision, manuscript discussion, and writing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors are grateful to the Conselho Nacional de Desenvolvimento Tecnológico (CNPq), Fundação Araucária – Programa de Pesquisa para o SUS (PPSUS grant 2016), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) for providing the financial support of the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01698/full#supplementary-material

REFERENCES

1. Choudri BS, Charabi Y, Ahmed M. Pesticides and Herbicides. *Water Environ Res.* (2018) 1:90.
2. Kim KH, Kabir E, Jahan SA. Exposure to pesticides and the associated human health effects. *Sci Total Environ.* (2017) 1:575. doi: 10.1016/j.scitotenv.2016.09.009
3. He TT, Zuo AJ, Wang JG, Zhao P. Organochlorine pesticides accumulation and breast cancer: a hospital-based case-control study. *Tumor Biol.* (2017) 1:6. doi: 10.1177/1010428317699114
4. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, et al. Risk factors and preventions of breast cancer. *Int J BiolSci.* (2017) 13:11. doi: 10.7150/ijbs.21635
5. Gearhart-Serna LM, Hoffman K, Devi GR. Environmental quality and invasive breast cancer. *Cancer Epidemiol Biomarkers Prev.* (2020). doi: 10.1158/1055-9965.EPI-19-1497 [Epub ahead of print].
6. Silva AMC, Campos PHN, Mattos IE, Hajat S, Lacerda EM, Ferreira MJM. Environmental exposure to pesticides and breast cancer in a region of intensive agribusiness activity in brazil: a case-control study. *Int J Environ Res Public Health.* (2019) 16:20. doi: 10.3390/ijerph16203951

7. Alleva R, Manzella N, Gaetani S, Bacchetti T, Bracci M, Ciarapica V, et al. Mechanism underlying the effect of long-term exposure to low dose of pesticides on DNA integrity. *Environ Toxicol.* (2018) 33:22534. doi: 10.1002/tox.22534
8. Bradlow HL, Davis DL, Lin G, Sepkovic D, Tiwari R. Effects of pesticides on the ratio of 16 alpha/2-hydroxyestrone: a biologic marker of breast cancer risk. *Environ Health Perspect.* (1995) 103:7. doi: 10.1289/ehp.95103s7147
9. Pires BRB, Panis C, Alves VD, Herrera ACSA, Binato R, Pizzatti L, et al. Label-free proteomics revealed oxidative stress and inflammation as factors that enhance chemoresistance in luminal breast cancer. *Oxid Med Cell Longev.* (2019) 2019:15. doi: 10.1155/2019/5357649
10. Corrêa S, Panis C, Binato R, Herrera AC, Pizzatti L, Abdelhay E. Identifying potential markers in breast cancer subtypes using plasma label-free proteomics. *J Proteomics.* (2017) 16:15. doi: 10.1016/j.jprot.2016.07.030
11. Panis C, Pizzatti L, Herrera AC, Corrêa S, Binato R, Abdelhay E. Label-free proteomic analysis of breast cancer molecular subtypes. *J Proteome Res.* (2014) 13:11. doi: 10.1021/pr500676x
12. Panis C, Pizzatti L, Corrêa S, Binato R, Lemos GF, Da Silva do Amaral Herrera AC, et al. The positive is inside the negative: HER2-negative tumors can express the HER2 intracellular domain and present a HER2-positive phenotype. *Cancer Lett.* (2015) 357:1. doi: 10.1016/j.canlet.2014.11.029
13. Pizzatti L, Panis C, Lemos G, Rocha M, Cecchini R, Souza GHMF, et al. Label-free MSE proteomic analysis of chronic myeloid leukemia bone marrow plasma: disclosing new insights from therapy resistance. *Proteomics.* (2012) 12:2618–31. doi: 10.1002/pmic.201200066
14. Velásquez E, Nogueira FCS, Velásquez I, Schmitt A, Falkai P, Domont GB, et al. Synaptosomal proteome of the orbitofrontal cortex from schizophrenia patients using quantitative label-free and iTRAQ-based shotgun proteomics. *J Proteome Res.* (2017) 16:4481–94. doi: 10.1021/acs.jproteome.7b00422
15. Gjertsen BT and Wiig H. Investigation of therapy resistance mechanisms in myeloid leukemia by protein profiling of bone marrow extracellular fluid. *Expert Rev Proteomics.* (2012) 9:595–8. doi: 10.1586/epr.12.55
16. Panis C, Pizzatti L, Herrera AC, Cecchini R, Abdelhay E. Putative circulating markers of the early and advanced stages of breast cancer identified by high-resolution label-free proteomics. *Cancer Lett.* (2013) 330:1. doi: 10.1016/j.canlet.2012.11.020
17. Ma Y, Ren Y, Dai ZJ, Wu CJ, Ji YH, Xu J. IL-6, IL-8 and TNF-α levels correlate with disease stage in breast cancer patients. *Adv Clin Exp Med.* (2017) 26:3. doi: 10.17219/acem/62120
18. Panis C, Pizzatti L, Souza GF, Abdelhay E. Clinical proteomics in cancer: where we are. *Cancer Lett.* (2016) 382:231–9.
19. Wetmore BA, Merrick BA. Toxicoproteomics: proteomics applied to toxicology and pathology. *Toxicol Pathol.* (2004) 32:6. doi: 10.1080/01926230490518244
20. Gonzalez H, Hagerling C, Werb Z. Roles of the immune system in cancer: from tumor initiation to metastatic progression. *Genes Dev.* (2018) 32:19–20. doi: 10.1101/gad.314617.118
21. Friedrich K. Desafios para a avaliação toxicológica de agrotóxicos no Brasil: desregulação endócrina e imunotoxicidade. *Vigilância Sanitária Em Debate.* (2013) 1:30. doi: 10.3395/vd.v1i2.30
22. Pardo LA, Freeman LEB, Lerro CC, Andreotti G, Hofmann JN, Parks CG, et al. Pesticide exposure and risk of aggressive prostate cancer among private pesticide applicators. *Environ Health.* (2020) 19:30. doi: 10.1186/s12940-020-00583-0
23. Volker D, Wolfgang H, Klausdieter B, Caner S, Christian C, Gerhard O. Associations of blood levels of PCB, HCHs and HCB with numbers of lymphocyte subpopulations, in vitro lymphocyte response, plasma cytokine levels, and immunoglobulin autoantibodies. *Environ Health Perspect.* (2001) 109:173. doi: 10.1289/ehp.01109173
24. Ruiz-Guzmán JA, Gómez-Corrales P, Cruz-Esquivel Á, Marrugo-Negrete JL. Cytogenetic damage in peripheral blood lymphocytes of children exposed to pesticides in agricultural areas of the department of Cordoba, Colombia, mutat res–genet toxicol environ mutagen. *Mutat Res.* (2017) 824:25–31. doi: 10.1016/j.mrgentox.2017.10.002
25. Volker D, Wolfgang H, Klausdieter B, Caner S, Christian C, Gerhard O. Associations of Dichlorodiphenyltrichloroethane (DDT) 4.4 and dichlorodiphenyldichloroethylene (DDE) 4.4 blood levels with plasma IL-4. *Arch Environ Health.* (2002) 57:6. doi: 10.1080/00039890209602086
26. Geier A, Dietrich CG, Voigt S, Kim SK, Gerloff T, Kullak-Ublick GA, et al. Effects of proinflammatory cytokines on rat organic anion transporters during toxic liver injury and cholestasis. *Hepatology.* (2003) 38:2. doi: 10.1053/jhep.2003.50317
27. Sifakis S, Androutsopoulos VP, Tsatsakis AM, Spandidos DA. Human exposure to endocrine disrupting chemicals: effects on the male and female reproductive systems. *EnvironToxicol Pharmacol.* (2017) 51:24. doi: 10.1016/j.etap.2017.02.024
28. McAnena PF, Brown JAL, Ramli A, Curran C, Malone C, McLaughlin R, et al. Breast cancer subtype discordance: impact on post-recurrence survival and potential treatment options. *BMC Cancer.* (2018) 18:203. doi: 10.1186/s12885-018-4101-7
29. Shiino S, Kinoshita T, Yoshida M, Jimbo K, Asaga S, Takayama S. Prognostic impact of discordance in hormone receptor status between primary and recurrent sites in patients with recurrent breast cancer. *Clin Breast Cancer.* (2016) 16:4. doi: 10.1016/j.clbc.2016.05.014
30. Lélu K, Laffont S, Delpy L, Paulet P-E, Perinat T, Tschanz SA, et al. Estrogen receptor α signaling in T lymphocytes is required for estradiol-mediated inhibition of Th1 and Th17 cell differentiation and protection against experimental autoimmune encephalomyelitis. *J Immunol.* (2011) 187:1578. doi: 10.4049/jimmunol.1101578
31. Lambert KC, Curran EM, Judy BM, Lubahn DB, Estes DM. Estrogen receptor-α deficiency promotes increased TNF-α secretion and bacterial killing by murine macrophages in response to microbial stimuli in vitro. *J Leukoc Biol.* (2004) 75:1166–72. doi: 10.1189/jlb.1103589
32. Gameiro CM, Romão F, Castelo-Branco C. Menopause and aging: changes in the immune system – a review. *Maturitas.* (2010) 67:316–20. doi: 10.1016/j.maturitas.2010.08.003
33. Wang C, Christin JR, Oktay MH. Guo W lineage-biased stem cells maintain estrogen receptor positive and negative mouse mammary luminal lineages. *Cell Rep.* (2017) 18:71. doi: 10.1016/j.celrep.2017.02.071
34. kłuciński P, Kossmann S, Friedek D. Release of interleukin-1 beta from peripheral blood monocytes and its concentration in the blood serum of the workers employed at the production of chlorfenvinphos. *Przegl Lek.* (1997) 54:10.
35. Chan DS, Norat T. Obesity and breastcancer: not only a risk factor of the disease. *Curr Treat Options Oncol.* (2015) 16:5. doi: 10.1007/s11864-015-0341-9
36. Ceelie H, Spaargaren-Van Riel CC, De Jong M, Bertina RM, Vos HL. Functional characterization of transcription factor binding sites for HNF1-alpha, HNF3-beta (FOXA2), HNF4-alpha, Sp1 and Sp3 in the human prothrombin gene enhancer. *J Thromb Haemost.* (2003) 1:8. doi: 10.1046/j.1538-7836.2003.00393.x

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 © 2020 Pizzatti, Kawassaki, Fadel, Nogueira, Evaristo, Woldmar, Teixeira, Da Silva, Scandolara, Rech, Candiotto, Silveira, Pavanelli and Panis. 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: 31 August 2020
doi: 10.3389/fcell.2020.00843

FBXO2 Promotes Proliferation of Endometrial Cancer by Ubiquitin-Mediated Degradation of FBN1 in the Regulation of the Cell Cycle and the Autophagy Pathway

Xiaoxia Che[1,2†], Fangfang Jian[1†], Ying Wang[3], Jingjing Zhang[1], Jian Shen[1], Qi Cheng[2], Xi Wang[2], Nan Jia[2] and Weiwei Feng[1]*

[1] Department of Obstetrics and Gynecology, Ruijin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [2] Obstetrics and Gynecology Hospital, Fudan University, Shanghai, China, [3] Department of Obstetrics and Gynecology, Yidu Central Hospital of Weifang, Weifang, China

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Maria Francesca Baietti,
VIB KU Leuven Center for Cancer Biology, Belgium
Jessica Dal Col,
University of Salerno, Italy

***Correspondence:**
Weiwei Feng
Fww12066@rjh.com.cn

[†] These authors have contributed equally to this work

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

Received: 01 June 2020
Accepted: 06 August 2020
Published: 31 August 2020

Citation:
Che X, Jian F, Wang Y, Zhang J, Shen J, Cheng Q, Wang X, Jia N and Feng W (2020) FBXO2 Promotes Proliferation of Endometrial Cancer by Ubiquitin-Mediated Degradation of FBN1 in the Regulation of the Cell Cycle and the Autophagy Pathway. Front. Cell Dev. Biol. 8:843. doi: 10.3389/fcell.2020.00843

F-box proteins, as substrates for S phase kinase-associated protein 1 (SKP1)-cullin 1 (CUL1)-F-box protein (SCF) ubiquitin ligase complexes, mediate the degradation of a large number of regulatory proteins involved in cancer processes. In this study, we found that F-box only protein 2 (FBXO2) was up-regulated in 21 endometrial carcinoma (EC) samples compared with five normal endometrium samples based on our Fudan cohort RNA-sequencing. The increased FBXO2 expression was associated with tumor stage, tumor grade, and histologic tumor type, and poor prognosis based on The Cancer Genome Atlas (TCGA) database. FBXO2 knockdown inhibited EC cell proliferation, and FBXO2 overexpression promoted the parental cell phenotype *in vivo* and *in vitro*. Fibrillin1 (FBN1) was also identified as a substrate for FBXO2 using a ubiquitination-proteome approach. In addition, promotion of EC proliferation by FBXO2 was regulated by specific proteins of the cell cycle (CDK4, CyclinD1, CyclinD2, and CyclinA1) and the autophagy signaling pathway (ATG4A and ATG4D) based on RNA sequencing (RNA-seq). We concluded that FBXO2 acts as an E3 ligase that targets FBN1 for ubiquitin-dependent degradation, so as to promote EC proliferation by regulating the cell cycle and the autophagy signaling pathway. Targeting FBXO2 may represent a potential therapeutic target for EC.

Keywords: endometrial carcinoma, FBXO2, FBN1, ubiquitination, cell cycle, autophagy

INTRODUCTION

Endometrial carcinoma (EC) is the commonest gynecologic malignancy in developed countries and the fourth most common cancer in women worldwide, and its prevalence continues to increase (Morice et al., 2016). It is estimated that there will be 61,880 new cases and 12,160 deaths in the United States from EC in 2019 (Siegel et al., 2019). Therefore, it is of paramount importance that we seek new biomarkers to improve the health and survival of patients with EC.

Ubiquitination by the ubiquitin-proteasome system (UPS) is a post-translational modification that regulates various cellular processes. It is carried out by a three-step cascade of ubiquitin transfer reactions (activation, conjugation, and ligation) regulated by ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligases (E3), respectively (Vucic et al., 2011). F-box proteins are the substrate adaptors of S phase kinase-associated protein 1 (SKP1)-cullin 1 (CUL1)-F-box protein (SCF) ubiquitin-ligase complexes that mediate the degradation of a number of regulatory proteins involved in cancer processes including cell growth, cell division, signaling responses, and cell survival and death (Skaar et al., 2013). F-box proteins can be classified into three families based on the presence of recognizable domains beyond the F-box domain (Cenciarelli et al., 1999; Ilyin et al., 2002). Members of this family have been shown to be essential in the regulation of cell proliferation and to exhibit oncogenic or tumor-suppressive activities or in cancer drug resistance (Uddin et al., 2016; Yan et al., 2019; Liu et al., 2020). Inactivation of FBXW7 resulted in the formation of precancerous lesions and well-differentiated endometrioid adenocarcinomas (Cuevas et al., 2019). High expression of SKP2 has been reported to be correlated with poor prognosis in endometrial endometrioid adenocarcinoma (Kamata et al., 2005).

F-box only protein 2 (FBXO2) also named Fbx2, Fbs1, and NFB42 is a cytoplasmic protein and ubiquitin ligase F-box protein with specificity for high-mannose glycoproteins (Yoshida et al., 2002). FBXO2 also was reported to act as a posttranscriptional regulator of hepatic insulin signaling and might constitute a novel therapeutic target for treating metabolic disorders (Liu et al., 2017). FBXO2 was found to regulate the EMT signaling pathway in gastric cancer and colorectal cancer suggesting that its importance involved in cancers (Sun et al., 2018; Wei et al., 2018). These investigators showed that F-box proteins play an essential role in cancers, and may provide potential strategies with which to treat them. However, whether FBXO2 plays a role in EC remains poorly understood.

In this study, we report our novel findings on the important role of FBXO2 in EC progression and investigate the underlying mechanisms of its action. We hypothesize that FBXO2 can serve as a potential therapeutic target for the precise treatment of EC.

MATERIALS AND METHODS

Clinical Specimens

This present study was approved by the ethics committee of Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine and Obstetrics and Gynecology Hospital of Fudan University. Twenty-one EC samples (three with clear cell carcinoma, five with serious carcinoma, and 13 with endometrioid endometrial carcinoma) and five normal samples (all of proliferative endometrium) from the Obstetrics and Gynecology Hospital, Fudan University (the "Fudan Cohort") were used to perform RNA sequencing (RNA-seq). Tissue microarrays containing 90 paired formalin-fixed and paraffin-embedded specimens (endometrioid endometrial carcinoma and para-carcinoma tissues) from tissue representing different stages of EC in patients who underwent surgical resection at Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine between 2013 and 2017 were examined. Among those samples, we extracted RNA from 19 normal endometrium and 47 EC tissues for subsequent validation by qRT-PCR. We also extracted proteins from 6 normal endometrium and six EC tissues for western immunoblotting analysis.

EC Cell Lines and Regents

AN3-CA, HEC-1B, ECC-1, KLE, and Ishikawa were kindly provided by Jie Jiang professor, Qilu Hospital of Shandong University. HEK-293T was kindly provided by Yingjie Xu professor, Ruijin Hospital of Shanghai Jiao Tong University. HEC-1A and RL95-2 cells were obtained from Chinese Academy of Science[1]. AN3-CA, HEC-1B, ECC-1, and KLE were maintained in RPMI-1640, Ishikawa and HEK-293T in DMEM, HEC-1A in McCoy's 5A, RL95-2 in F12 medium supplemented with 10% fetal bovine serum (FBS, Gibco Inc., Albany, NY, United States), 100 μg/ml penicillin/streptomycin and 2 mM L-glutamine in a humidified incubator at 37°C and 5% CO2. All cells tested negative for mycoplasma contamination and the certificate of STR analysis were included for all cell lines. 3-Methyladenine (3-MA) (Sigma-Aldrich, St. Louis, MO, United States), palbociclib (MedChemExpress, Trenton, NJ, United States), MG132 (Selleckchem, Houston, TX, United States, S2619), and cycloheximide (CHX) (Enzo, ALX-380269) were purchased from the mentioned manufactures.

Establishment of FBXO2 Stable Knockdown and Overexpressing Cells, and FBN1 Stable Knockdown Cells and FBN1 Transient Transfection

For the construction of stable FBXO2 knock down cell line, we first selected three interfering plasmids to test their efficiency both in mRNA and protein levels shown in **Figure 2A**. And we selected the shFBXO2-1 and shFBXO2-2 to establish stable cell lines. We used two shRNAs (shFBXO2-1 and shFBXO2-2) in the cells phenotype (named RL95-2-shFBXO2-1 and RL95-2-shFBXO2-2 in **Figure 2**), in the following mechanism and *in vivo* sections, we chose the most effective shRNA (shFBXO2-2) for further study (named RL95-2-shFBXO2 in **Figures 3–7**). The FBXO2, FBN1, and FBXO2/FBN1 stable knockdown cell lines (RL95-2-shFBXO2-1, RL95-2-shFBXO2-2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1) were established by lentiviral-based stable shRNA subcloned into the RNAi pLenti hU6-MCS-CMV-zsGreen1-PGK-Puro vector (LncBio Co., Shanghai, China) (shFBXO2-1 target sequence: TGGTGTGACGTGGAGCATGGT; shFBXO2-2 target sequence: GGAGTTCACCCACGATGAGAG; shFBXO2-3 target sequence: TCGTGGTGAAGGACTGGTACT; shFBN1 target sequence: CAGCTGGCATCAGATGGACGTTATT). Non-target control shRNA served as a negative control (RL95-2-NC). The FBXO2 stably overexpressing cell line

[1] www.cellbank.org.cn

(Ishikawa-ovFBXO2) was established by lentiviral-based stable LV-FBXO2 subcloned into the GV492 Ubi-MCS-3FLAG-CBh-gcGFP-IRES-puromycin vector (GeneChemBio Co., Shanghai, China). The transient overexpressing of FBN1 cell lines (RL95-2-ovFBN1 and Ishikawa-ovFBN1) was conducted using FBN1 plasmids cloned into a pcDNA3.1 vector (Target sequence: GAACAAAAACTCATCTCAGAAGAGGATCTG).

Plasmid Construction, Transfection, and Immunoprecipitation

Flag-tagged wild-type (WT), truncated, and mutant (MUT) FBXO2 (NM_012168), and Myc-tagged FBN1 (NM_000138) were subcloned into the pcDNA3.1 vector (LncBio Co., Shanghai, China). HEK-293T cells were transfected with the plasmids using Lipofectamine 3000 (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's instructions for FBXO2 and FBN1 binding analysis. We incubated protein A/G agarose with antibodies (5–10 μg) for 20 min, lysed the cells, and incubated the supernatants with the protein A/G agarose and antibodies for 1 h at room temperature. The incubation was boiled with 1 × Laemmli buffer for 10 min at 99°C. The immune complexes were subjected to SDS-PAGE and analyzed by immunoblotting. To immunoprecipitate the endogenous proteins, cells were lysed with 1 × cell lysis buffer. The supernatant was precleared with protein A/G agarose with indicated antibody overnight (FBXO2, Santa Cruz Biotechnology, sc-393873; FBN1, LifeSpan BioSciences, LS-B5512). Thereafter, the incubation was boiled with 1 × Laemmli buffer and analyzed by SDS-PAGE.

Subcutaneous Tumor Implantation Model

In total, 5×10^6 (0.1 ml) RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 cells were injected subcutaneously into 4-week-old immune-deficient BALB/c-nu mice (Shanghai LC Company). Tumor formation and mice weight was measured every 3 days, and tumor volume was calculated as 1/2 × length × width^2 for almost 4 weeks. All procedures were approved by the Animal Ethic Review Committee of Shanghai Jiao Tong University School of Medicine. All animals were handled according to the Guide for the Care and Use of Laboratory Animals' and the Principles for the Utilization and Care of Vertebrate Animals.

Reverse Transcription-Quantitative Polymerase Chain Reaction (qRT-PCR)

Total RNA was extracted from EC cell lines and patients tissue via TRIzol (Thermo Fisher Scientific Inc., MA, United States) according to the manufacturer's instrument. Reverse transcribe mRNA to cDNA and real-time PCR were performed by the PrimeScript™ Reagent Kit and SYBR Premix Ex Taq™ II (Takara Bio Inc., Shiga, Japan). The primers were synthesized from BioTNT (Shanghai BioTNT Co., Ltd.). The primers sequence of the following genes are described in **Table 1**.

TABLE 1 | The primers sequence in the article.

Primer	Sequence (5′ to 3′)
FBXO2-F	ACTTGGAAGGCTGGTGTGAC
FBXO2-R	TCAAAGGAGGAGGCGAAGTA
FBN1-F	TGCCGCATATCTCCTGACCTCTG
FBN1-R	TAGCCTTCGTCACACTTGCATTCG
GAPDH-F	GGGAAGGTGAAGGTCGGAGT
GAPDH-R	GGGGTCATTGATGGCAACA
CCND1-F	GAGACCATCCCCCTGACGGC
CCND1-R	TCTTCCTCCTCCTCGGCGGC
CCND2-F	CATCCTCACGGCCTCGGCT
CCND2-R	CGGCGTGTGTTTATCGGAATCCA
THBS1-F	AGACTCCGCATCGCAAAGG
THBS1-R	TCACCACGTTGTTGTCAAGGG
CHEK1-F	TCTTTGGACTCGCTCAAGAAGCCT
CHEK1-R	ATTTCAACCTTCGGTGTGCTTGGG
CDK4-F	GAAGAAG AAGCGGAGGAAGAGG
CDK4-R	TTAGGTTAGTGCGGGAATGAAT
SMC1A-F	AAAGCATCAAGCGCCTTTAC
SMC1A-R	GCTGGCATAGGTCAATGAGG
MCM7-F	CCTACCAGCCGATCCAGTCT
MCM7-R	CCTCCTGAGCGGTTGGTTT
CDC14B-F	CAAACGCTTTACGGATGCTGG
CDC14B-R	TGATGTAGCAGGCTATCAGAGT
PPP2CA-F	GATCTTCTGTCTACATGGTGGTCTC
PPP2CA-R	ACACATTGGACCCTCATGGGGAA
ATG4A-F	TTGGCCCAGGATGACAGCTG
ATG4A-R	AGGGCCCGTTCCACCAATTG
ATG4B-F	TGAGTCTTGTGGTGTGTGGT
ATG4B-R	TACTTTCCCAGGACAGGCAG
ATG4C-F	GTTACCTGCAGAGTCGGGAT
ATG4C-R	GGCCAGTTCTCAATGTGCAG
ATG4D-F	GTCCATGAACTCAGTGTCGC
ATG4D-R	GAACTTGTCCACTTCGTCCG

Immunofluorescence (IF) Staining

Ishikawa and RL95-2 cells were incubated with anti-FBXO2 (1:50, Santa Cruz Biotechnology, sc-398111) and anti-FBN1 antibodies (1:100, Thermo Fisher Scientific, PA5-82743) in PBS at 4°C overnight, and then with Alexa Fluor 488-conjugated and 555-conjugated donkey anti-goat secondary antibody (1:1000, Abcam; 1:1000, Life Technology). Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) (Sigma, D9542) for 10 min. Confocal microscopy (Leica, Germany) was used to capture images.

Tissue Microarray Assay Staining

Tissue microarrays containing 90 paired formalin-fixed and paraffin-embedded specimens of EC were obtained from Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine. Slides of patients tissue were incubated with anti-FBXO2 rabbit polyclonal (1:100, Abcam, ab230307), anti-FBN1 rabbit polyclonal (1:50, Arigobio, ARG66654), and Ki67 (1:100, GB13030-2, Goodbiotechnology, Wuhan, China) antibodies. We assessed the expression levels of FBXO2 and FBN1 by the

percentage of positively stained cells, and the IHC index was determined on the basis of the staining intensity and the percentage of immuno-reactive cells. Staining intensity was rated as 0 (negative), 1 (weakly positive), 2 (moderately positive), or 3 (strongly positive), and the percentage of staining was rated as 0 (0%), 1 (1-25%), 2 (26-50%), 3 (51-75%), or 4 (76-100%). The IHS (immunohistochemistry score) was calculated by multiplying the quantity and intensity scores ranging from 0 to 12. An IHS score of 0–6 was low and 7–12 was high immunoreactivity as previously described (Zhi et al., 2010). The expression of FBXO2, FBN1, and Ki67 in mouse xenografts was quantified using Image Pro Plus (IPP), involving the three parameters of integrated optical density (IOD), area, and mean optical density were involved: IOD sum/area. Images were obtained with a microscope (Olympus BX53; Olympus, Tokyo, Japan) fitted with a digital camera (Olympus DP73; Olympus). Five randomly selected images at ×400 magnification of each sample were taken to achieve a mean optional density value with Image Pro-Plus 6.0 (version 6.0.0.206; Media Cybernetics, Inc., Bethesda, MD, United States).

Flow Cytometry for Analysis of Cell Cycle and Apoptotic Cells

To analyze the cell cycle, cells were seeded in six-well culture plates and cultured to 90% confluence and were harvested and permeabilized with cold 75% ethanol at −20°C overnight. Then the cells were centrifuged at 150 × *g* for 10 min at room temperature, the precipitate was resuspended in 2 ml of 0.9% physiologic saline and centrifuged at 150 × *g* for 10 min. Cell cycle progression was evaluated after propidium iodide (PI) staining for 30 min in the dark.

We identified apoptotic cells using the Annexin V/Dead Cell Apoptosis Kit (BD-Pharmingen). Briefly, cells were seeded in 6-well plates for 48 h to 95% confluence. Then cells were harvested and washed twice with ice-cold PBS resuspended in 1× annexin-binding buffer and incubated with 5 μl of Annexin V-PE (phycoerythrin) and 5 μl of 7-AAD (7-amino-actinomycin) at room temperature for 15 min in the dark. PE and 7-AAD fluorescence was analyzed by flow cytometry within 1 h.

Proliferation Potential Analysis

Colony Formation Assay

Ishikawa-NC, Ishikawa-ovFBXO2, RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 cells were seeded in six-well plates at 600, 5000 cells/well, respectively, and then incubated for 12 days as previously in Che et al. (2016).

Cell Counting Kit-8 (CCK-8) Assay

Cell proliferation was also measured by the CCK-8 assay kit (Dojindo Japan). Briefly, cells with 4000–10,000 density were seeded in 96-well culture plates for 6–7 days. After incubation for indicated time, 10 μl CCK-8 reagent was added to each well for incubating 1 h and the absorbance was measured at 450 nm.

A EdU Assay

A EdU assay (Riobio, Guangzhou, China) was used in proliferation assays performed in a 96-well format with 4,000 and 10,000 cells of Ishikawa and RL95-2 cell lines, respectively, according to the manufacturer's instrument. Images were captured by fluorescence microscope.

A Real-Time Cell Proliferation Assay (RTCA)

Proliferation assays were also performed in 16-well CIM plates in an xCELLigence DP instrument for real-time cell proliferation analysis (Roche, Mannheim, Germany). Cells (Ishikawa: 4,000 cells/well, and RL95-2: 10,000 cells/well) were seeded in the plate and grow on top of electrodes so that the impedance varies on the basis of the number of cells attached and the quality of cell-electrode interaction. Electrode impedance, displayed as cell index, was used to monitor cell viability.

RNA Sequencing Analysis

The RNA expression profile was determined in the sequencing libraries generated form a NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, United States). The clustering of the samples were performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina), and the library preparations were sequenced on an Illumina NovaSeq platform. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, and Gene Set Enrichment Analysis (GSEA) were performed to analyze the FBXO2 regulated genes based on the sequencing data sets. The RNA-seq data was available in **Supplementary Data** named "21 EC VS 5 normal samples diff genes" and "shFBXO2 VS Negtive Control diff genes."

K-ε-GG Profiling and Proteome Analysis by Liquid Chromatograph-Tandem Mass Spectrometry (LC MS/MS)

The supernatant from each sample, containing precisely 10 mg of protein was digested with Trypsin Gold (Promega) at a 1:50 enzyme-to-substrate ratio. After 16 h of digestion at 37°C, peptides were desalted with a C18 cartridge to remove the high urea, and desalted peptides were dried by vacuum centrifugation. MOPS IAP buffer (50 mM MOPS, 10 mM KH_2PO_4, and 50 mM NaCl) was added to resuspend the lyophilized peptides. Then centrifuged peptides for 5 min at 12000 *g*. Supernatants were mixed with anti-Ubiquitin Remnant Motif K-ε-GG beads (CST #5562, Cell Signaling Technology) for 2.5 h at 4°C. Beads were washed in MOPS IAP buffer, and then in water, prior to elution of the peptides with 0.15% TFA. In preparation for analysis, the peptides were desalted using peptide desalting spin columns (Thermo Fisher, 89852).

We performed shotgun proteomics analyses using an EASY-nLC™ 1200 UHPLC system (Thermo Fisher) coupled with an Orbitrap Q Exactive HF-X mass spectrometer (Thermo Fisher) operating in the data-dependent acquisition (DDA) mode, and the raw data were processed using [the Proteome Discoverer 2.2 (PD) 2.2, Thermo] search engines. Tandem mass spectra were searched against the Swissprot human (20,274 sequences) database concatenated with a reverse decoy database. The K-ε-GG profiling and proteome analysis by liquid chromatograph-tandem mass spectrometry was performed in Novogene Co., Ltd.,

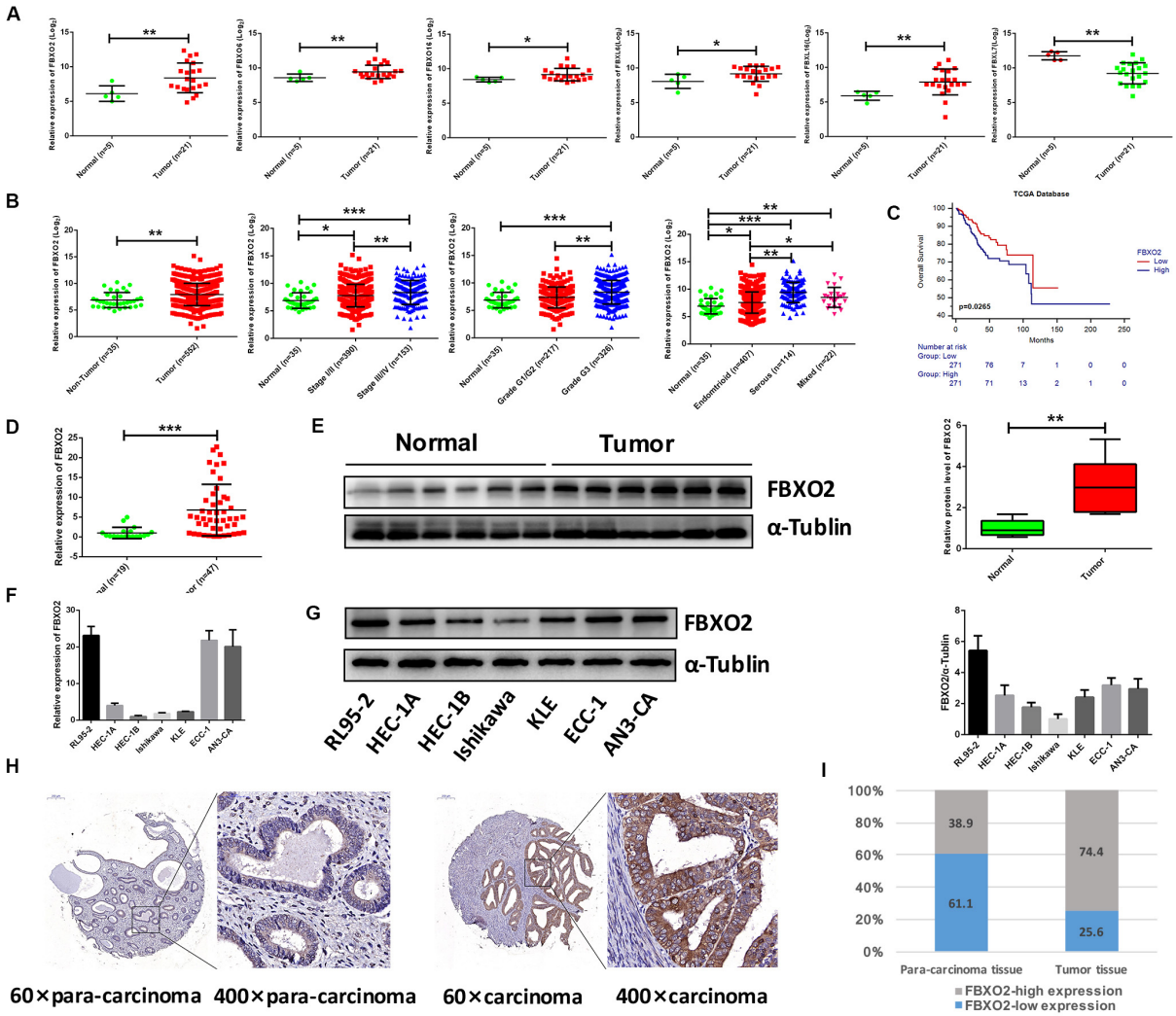


FIGURE 1 | Frequent dysregulation of the F-box family in human EC based on the RNA-sequencing data in the Fudan cohort, and the frequent up-regulation of FBXO2 in the TCGA database, and in human EC tissues. **(A)** The mRNA expression of F-box family members (FBXO2, FBXO6, FBXO16, FBXL6, FBXL16, and FBXL7) in our RNA-sequencing data (*$P < 0.05$, **$P < 0.01$, *t*-test). **(B)** FBXO2 mRNA expression in 552 EC samples compared with 35 non-tumorous samples, based on the TCGA database (**$P < 0.01$, *t*-test); FBXO2 mRNA expression at different stages, grades and histologic types of ECs (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$, one-way ANOVA). **(C)** The Kaplan-Meier survival analysis of EC patients with different expression levels of FBXO2 ($P = 0.0265$, log-rank test). **(D)** FBXO2 mRNA expression in 47 EC samples compared with 19 normal samples (***$P < 0.001$, *t*-test). **(E)** Western immunoblotting analysis of FBXO2 protein expression in six EC samples compared with six normal samples (**$P < 0.01$, *t*-test). **(F,G)** FBXO2 mRNA **(F)** and protein **(G)** expression in seven EC cell lines. **(H)** Immunohistochemical staining of FBXO2 in 90 EC samples and paired para-carcinoma samples. Representative photomicrographs from a patient. Original magnification, ×60 and ×400; scale bars: 200 and 50 μm, respectively. **(I)** Proportions of low or high expression in EC samples and para-carcinoma samples.

Beijing, China. The proteome sequencing data was available in **Supplementary Data** named "ishikawa ov VS Ishikawa nc diff proteins."

Statistical Analysis

Statistical analysis was performed using IBM SPSS 22.0 software (SPSS Inc., Chicago, IL, United States). The data all are presented as mean ± standard deviation (SD). Student's *t*-test and one-way analysis of variance were applied to assess the differences. The enumeration data was evaluated using χ^2 test or Fisher exact-probability test. The cumulative survival curves were drawn using the Kaplan-Meier method. $P < 0.05$ was considered to be statistically significant.

RESULTS

FBXO2 Is Up-Regulated in Endometrial Cancer and Is Correlated With Tumor Progression

To explore the mechanisms underlying the progression of EC, we performed RNA-seq analysis on normal endometrium

(n = 5) and EC samples (n = 21). A total of 918 genes were differentially expressed (641 up- and 277 down-regulated genes) in EC tissues compared with normal endometrium (with a padj < 0.01 and | Log2Fold change| > 2). F-box protein family genes were significantly enriched in our sequencing results, including FBXO2, FBXO6, FBXO16, FBXL6, FBXL16, and FBXL7 (**Figure 1A**), and we chose FBXO2 for subsequent experiments as it was the most significantly up-regulated gene.

Consistent with our data, FBXO2 mRNA was increased in the TCGA database that consisted of 552 patients with endometrial cancer (**Figure 1B**). When we then divided the EC samples into different components according to the EC stage, grade, and histologic type. The results showed that FBXO2 expression in stage III/IV samples and grade 3 was higher than that in stage I/II and grade 1/2, respectively (**Figure 1B**). We also found that the FBXO2 expression in the serous type was significantly higher than in the endometrioid type (**Figure 1B**), which implied that FBXO2 high expression was related to EC progression. Moreover, high level of FBXO2 expression was significantly associated with shorter survival (**Figure 1C**). FBXO2 was also up-regulated in EC samples compared with normal samples both in the mRNA and protein levels in Fudan cohort (**Figures 1D,E**). Furthermore, when we evaluated the expression of FBXO2 using tissue microarrays. We observed that the proportion of highly expressed FBXO2 in EC samples was significantly higher than that in para-carcinoma samples (**Figures 1H,I**, 74.4 vs. 38.9%, $P < 0.05$). Next, we investigated whether high FBXO2 expression was related to clinicopathologic features and found that high FBXO2 expression was significantly associated with muscular infiltration and Ki67 staining (**Table 2**). Our results suggested that high expression of FBXO2 in EC samples was correlated with endometrial cancer progression.

TABLE 2 | Clinicopathological characteristics of Test trail patients according to FBXO2 expression.

Characteristics	Patients		FBXO2 expression			
	No.	%	High	Low	χ^2	*P*-value [a]
All patients	90	100	67	23		
Age (years)					**1**.76	0.185
<60	51	56.7	36	16		
≥60	39	43.3	31	7		
Tumor size (cm)					0.387	0.534
<4	64	71.1	51	13		
≥4	26	28.9	16	10		
Lymphovascular invasion					0.658	0.417
Positive	8	8.89	5	3		
Negative	82	91.1	62	20		
Lymph node metastasis					2.574	0.109
Metastasis	32	35.6	27	5		
No metastasis	58	64.4	40	18		
FIGO stage					1.045	0.307
I/II	73	81.1	56	17		
III/IV	17	18.9	11	6		
Muscular infiltration					5.711	**0.017**
<1/2	70	77.8	48	22		
≥1/2	20	22.2	19	1		
Ki67 staining					7.809	**0.005**
<50%	52	57.8	33	19		
≥50%	38	42.2	34	4		

[a] *P-value < 0.05 marked in bold font shows statistical significant.*

FBXO2 Promotes Proliferation of EC Cells

To explore whether FBXO2 plays a crucial role in EC progression, we selected 2 EC cell lines (RL95-2 with FBXO2 high expression, and Ishikawa with FBXO2 low expression) (**Figures 1F,G**) to further validate the function of FBXO2. CCK-8, RTCA, colony formation assay, and EdU staining showed that FBXO2 knockdown significantly inhibited proliferation of RL95-2-shFBXO2-1/2 cells, while overexpression in Ishikawa-ovFBXO2 cells significantly promoted proliferation (**Figures 2D–G**). In addition, cell cycle analysis showed that FBXO2 knockdown in RL95-2-shFBXO2-1/2 cells increased the proportion of cells in G1 phase and decreased the proportion of cells in S phase; while FBXO2 overexpression in Ishikawa-ovFBXO2 cells decreased the proportion of cells in G1 phase and increased the proportion of cells in G2 phase (**Figures 2H,I**). Collectively, these results suggested that FBXO2 critically regulated EC cells proliferation.

FBXO2 Interacts With FBN1 and Negatively Regulates the Stability of the FBN1

To identify the proteome-wide changes in ubiquitylated proteins associated with the oncogenic tumor function of FBXO2, we isolated the FBXO2 protein complex from Ishikawa-ovFBXO2 and Ishikawa-NC stable cell lines mixed with anti-ubiquitin remnant motif, and determined the proteins present in the complex by using liquid chromatograph-tandem mass spectrometry (LC MS/MS) to interrogate the differences in the proteome and ubiquitylome between Ishikawa-NC and Ishikawa-ovFBXO2 cells. As verification of the efficiency of this approach, we obtained the peptides that were ubiquitin-mediated by FBXO2 as shown in **Figures 3A,B**. To obtain the FBXO2-interacting proteins, Liu et al. (2017) isolated the FBXO2 protein complex from HEK293T cells expressing Flag-tagged wild-type (WT) FBXO2 or F-box-associated domain mutant (MUT) that cannot recognize glycoproteins (Mizushima et al., 2007; Glenn et al., 2008), and determined the proteins by using high-performance liquid chromatography/tandem mass spectrometry analysis (IP-MS). The intersection of IP-MS, LC MS/MS, and Fudan cohort RNA-seq results revealed the fibrillin1 (FBN1) protein. Given that the interaction between FBXO2 and FBN1 has not been previously reported, we examined the potential functional relationship between them.

To clarify the interaction between FBXO2 and FBN1, we first measured endogenous FBN1 levels and the results showed that FBN1 mRNA levels were similar in RL95-2 and Ishikawa cells (**Figure 3E**), however, FBN1 protein levels were higher in

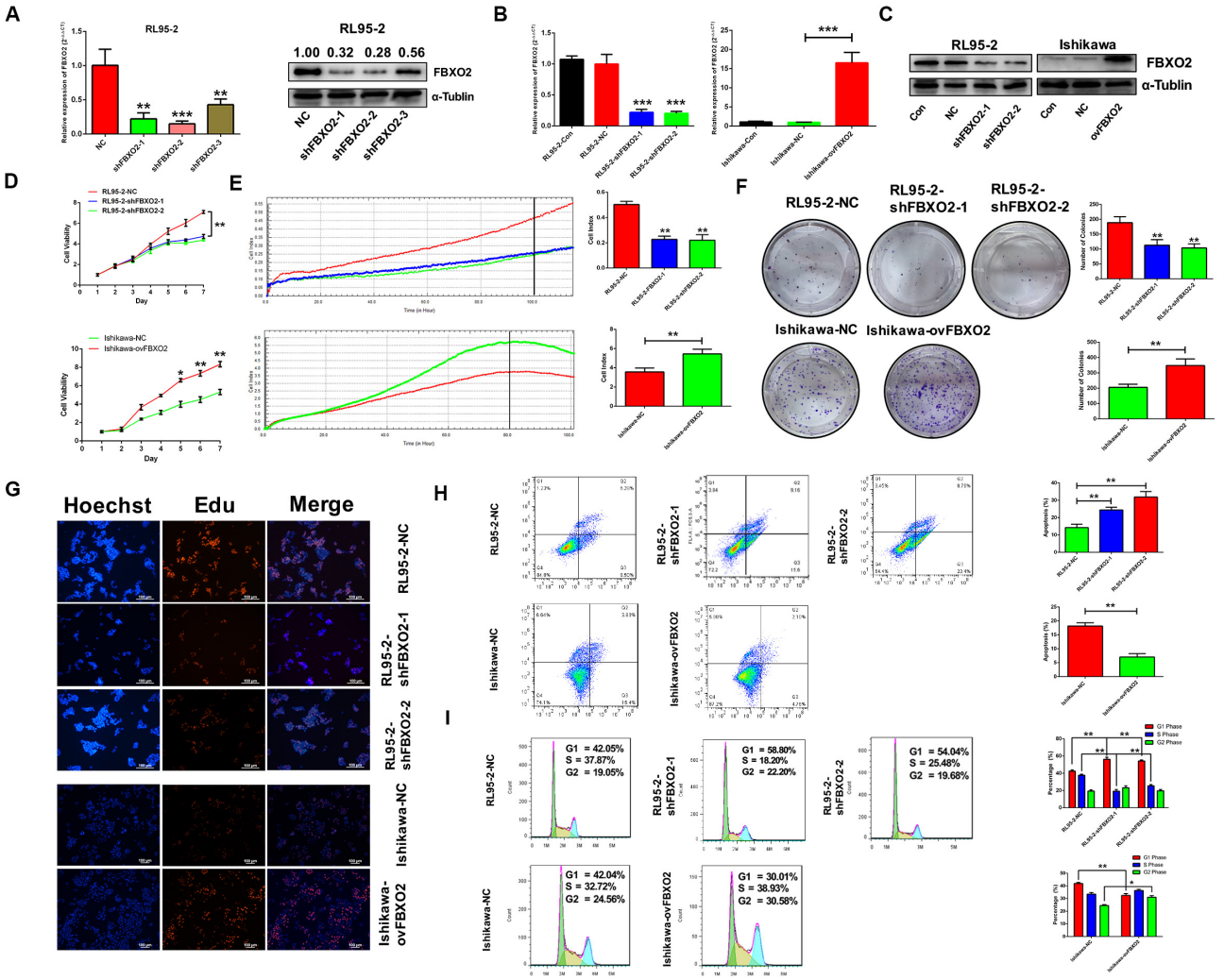


FIGURE 2 | FBXO2 promotes EC cells proliferation. **(A)** Three interfering plasmids were used to test FBXO2 interfering efficiency both in mRNA and protein levels in RL95-2 cell line. Data are presented as mean ± SD, $^{**}P < 0.01$, $^{***}P < 0.001$, *t*-test. **(B,C)** FBXO2 mRNA **(B)** and protein **(C)** levels in RL95-2-shFBXO2-1/2 and Ishikawa-ovFBXO2 cell lines compared with their parental cells. Data are presented as mean ± SD, $^{***}P < 0.001$, *t*-test. **(D–F)** EC cell proliferation was measured by CCK-8 assay **(D)**, RTCA **(E)**, and colony formation assay **(F)** in RL95-2-shFBXO2-1/2 and Ishikawa-ovFBXO2 cell lines compared with their parental cells. Data are presented as mean ± SD, $^{*}P < 0.05$, $^{**}P < 0.01$, *t*-test. **(G)** EdU staining in RL95-2-NC, RL95-2-shFBXO2-1/2, Ishikawa-NC, and Ishikawa-ovFBXO2 groups. **(H)** Apoptotic index was measured by flow cytometric assay in RL95-2-shFBXO2-1/2 and Ishikawa-ovFBXO2 cells compared with their parental cells. Data are presented as mean ± SD, $^{**}P < 0.01$, *t*-test. **(I)** Cell cycle analysis by flow cytometric assay in RL95-2-shFBXO2-1/2 and Ishikawa-ovFBXO2 cells compared with their parental cells. Data are presented as mean ± SD, $^{*}P < 0.05$ and $^{**}P < 0.01$, *t*-test.

Ishikawa cells and lower in RL95-2 cells which was inversely correlated with endogenous FBXO2 levels (**Figure 3F**). In addition, after knockdown of FBXO2, FBN1 protein levels were dramatically increased in RL95-2-shFBXO2 cells, while FBN1 protein was decreased after overexpression of FBXO2 in Ishikawa-ovFBXO2 cells (**Figure 3G**). Tissue array (IHC) results also showed that high expression of FBXO2 was correlated with low expression of FBN1 (**Figures 3H,I**, Pearson $r = -0.305$, $P = 0.004$). Furthermore, exogenous and endogenous Co-IP assays were performed to confirm the direct interaction between FBXO2 and FBN1 (**Figures 3J,K**). FBXO2 and FBN1 also co-localized in RL95-2 and Ishikawa cell cytoplasm and membrane as observed with confocal microscopy (**Figure 3L**).

Because FBXO2 interacted with FBN1, we tested whether FBXO2 regulated FBN1 stability or accelerated its protein degradation. The ubiquitination of FBN1 was also increased by ectopic expression of FBXO2 in HEK293T cells treated with MG132, a proteasome inhibitor (**Figure 4C**). Furthermore, overexpression of FBXO2 in Ishikawa-ovFBXO2 cells reduced the half-life of FBN1 compared with Ishikawa-NC cells while knockdown of FBXO2 in RL95-2-shFBXO2 cells augmented the half-life of FBN1 compared with RL95-2-NC cells (**Figures 4A,B**). These results supported the hypothesis that FBXO2 interacted with FBN1, and regulated its stability and promoted its degradation.

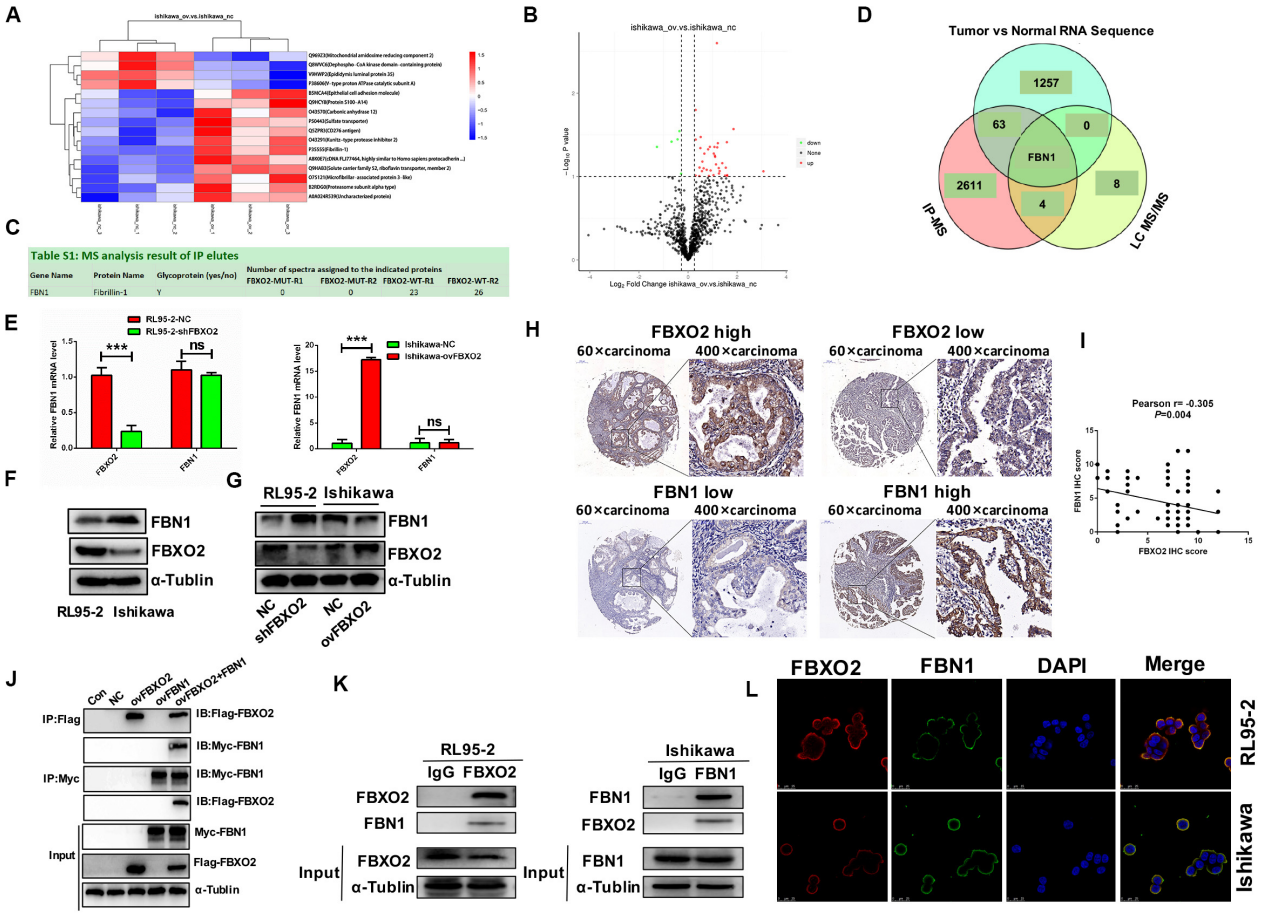


FIGURE 3 | FBN1 identified as a novel substrate for FBXO2. **(A,B)** Cluster analysis and volcano plot of the differential proteins regulated by FBXO2 ubiquitination based on the LC MS/MS data. **(C)** FBN1 protein was directly bound to FBXO2 protein. **(D)** The intersection of IP-MS, LC MS/MS, and our RNA sequencing data. **(E)** FBN1 mRNA expression measured by RT-PCR in RL95-2-shFBXO2 and Ishikawa-ovFBXO2 cells compared with their parental cells ($P > 0.05$). ***$P < 0.001$, *t*-test. **(F)** Endogenous FBXO2 and FBN1 protein levels in RL95-2 and Ishikawa cells. **(G)** FBN1 protein levels as measured by western blot analysis in shFBXO2-RL95-2 cells and Ishikawa-ovFBXO2 cells compared with their parental cells. **(H)** Immunohistochemical staining of FBXO2 and FBN1 in 90 EC samples and paired para-carcinoma samples. Left four photographs: FBXO2 was highly expressed in carcinoma tissues (upper panel), while FBN1 showed low expression in the same patient (lower panel). Right four photographs: FBXO2 showed low expression in carcinoma tissues (upper) while FBN1 was highly expressed in the same patient (lower). Original magnification, ×60 and ×400; scale bar: 200 and 50 μm, respectively. **(I)** Inverse correlation between FBXO2 and FBN1 staining, Pearson $r = -0.305$, $P = 0.004$. **(J)** Exogenous FBXO2 and FBN1 proteins interacted with each other in HEK-293T cells. HEK-293T cells were transfected with Flag-FBXO2, Myc-FBN1, and co-transfected with Flag-FBXO2 and Myc-FBN1 for 48 h, respectively. After treatment with 20 μM MG132 for 8 h, cell lysates were prepared for co-IP with anti-Flag or anti-Myc beads and western blot analysis. **(K)** Endogenous FBXO2 and FBN1 proteins interacted with each other in endometrial cancer cell lines. RL95-2 and Ishikawa cell lysates were prepared for co-IP with anti-FBXO2 or anti-FBN1 and western blot analysis. **(L)** FBXO2 and FBN1 co-localized in RL95-2 and Ishikawa cells cytoplasm and membrane. EC cells were immunostained with anti-FBXO2 (red) and anti-FBN1 (green) antibodies and visualized with confocal microscopy. DAPI (blue) was used to indicate cell nuclei. Scale bar, 25 μM.

As FBXO2 was proven to interact with FBN1 and promote its degradation by polyubiquitination, we next conducted experiments to determine which domain of FBXO2 serves this function. The F-box associated (FBA) domain of FBXO2 is essential for its glycoprotein-recognizing activity, and this activity is completely abolished by two residues mutations named Trp279A and Tyr278A (Mizushima et al., 2007; Glenn et al., 2008). Therefore, we constructed FBXO2 FL, FBXO2 N-terminal F-box domain, FBXO2 C-terminal FBA domain and their corresponding mutant structures (FBXO2-Tyr278A, FBXO2-Trp279A) were established (**Figure 4D**). The results showed that the FBXO2 C terminus down-regulated FBN1 protein levels in a dose-dependent manner while the FBXO2 N terminus did not exert this function (**Figure 4E**). We also constructed the FBXO2 Tyr278A and Trp279A mutant plasmids and results revealed that these two mutant plasmids did not down-regulate FBN1 protein levels (**Figure 4F**). To prove whether the FBXO2 FBA domain regulated FBN1 polyubiquitination levels, HA-ub and Myc-FBN1 were co-transfected into HEK-293T cells with various domains of FBXO2 and enzymatic dead mutant (Tyr278A and Trp279A). Our results revealed that the FL, and normal C-terminal domain of FBXO2 enhanced FBN1 polyubiquitination (**Figure 4G**). These results support the concept that FBXO2, along with FBA domain activity,

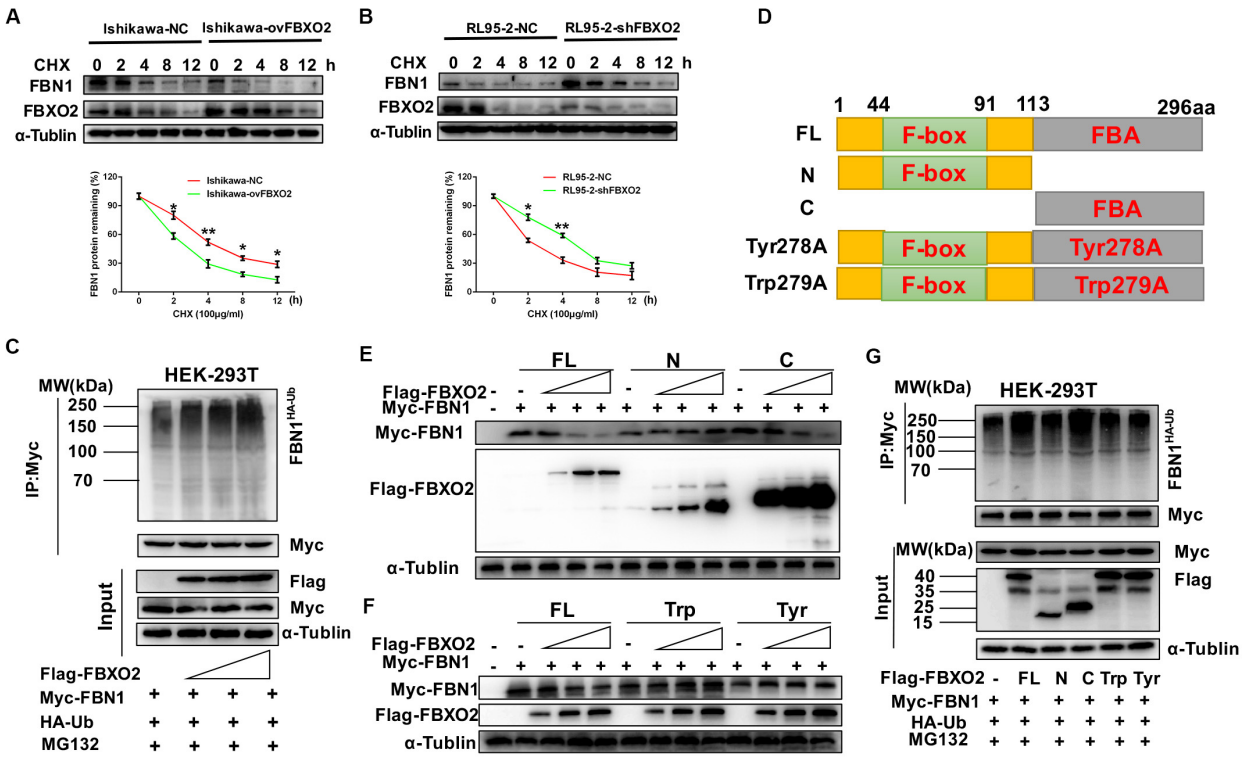


FIGURE 4 | FBXO2 degrades FBN1 protein through its ubiquitination. **(A,B)** Overexpression or knockdown of FBXO2 shortened or prolonged FBN1 protein half-life in Ishikawa or RL95-2 cells, respectively. Ishikawa-NC, Ishikawa-ovFBXO2, RL95-2-NC, and RL95-2-shFBXO2 cells were treated with 100 μg/ml of the protein synthesis inhibitor cycloheximide (CHX), and harvested at indicated time points for evaluation of FBN1 protein levels. $^{*}P < 0.05$, $^{**}P < 0.01$. **(C)** FBXO2 enhanced FBN1 polyubiquitination in a dose-dependent manner. HEK-293T cells were transfected with Flag-FBXO2, Myc-FBN1, and HA-ub for 48 h and then treated with 20 μM MG132 for 8 h. Cell lysates were prepared for co-IP with anti-Myc beads and western blotting analysis was used to detect the HA-ub (FBN1 ubiquitination levels). **(D)** Schematic representation of the FBXO2 protein. FL, full long; N, N terminus, including F-box domain; N, N terminus, including FBA domain; Tyr278A, FBA domain was replaced with non-functional Tyr278A mutant; Trp279A, FBA domain was replaced with non-functional Trp279A mutant. **(E)** The FBXO2 C terminus down-regulated FBN1 protein levels in a dose-dependent manner while the FBXO2 N terminus did not have the function. HEK-293T cells were transfected with 1.5 μg of Myc-FBN1 and Flag-FBXO2 FL/N/C terminus plasmids at doses of 0, 0.75, 1.5, or 3.0 μg for 48 h before harvesting for western blotting analysis. α-Tublin was used as a loading control. **(F)** The FBXO2 Tyr278A and Trp279A mutants did not down-regulate the FBN1 protein levels. HEK-293T cells were transfected with 1.5 μg Myc-FBN1 and Flag-FBXO2 FL/Tyr278A/Trp279A plasmids at doses of 0, 0.75, 1.5, or 3.0 μg for 48 h before harvesting for western blotting analysis. α-Tublin was used as a loading control. **(G)** The FBXO2 C terminus (FBA domain) enhanced FBN1 polyubiquitination. HEK-293T cells were transfected with Flag-FL/N/C/Tyr278A/Trp279A, Myc-FBN1, and HA-ub for 48 h and then treated with 20 μM MG132 for 8 h. Cell lysates were prepared for co-IP with anti-Myc beads and western blotting analysis was used to detect the HA-ub (FBN1 ubiquitination levels).

promoted FBN1 protein degradation through polyubiquitination of FBN1 protein.

FBN1 Is Required for FBXO2 to Exert Functional Impacts on EC Proliferation

To investigate whether FBN1 was required for FBXO2 to regulate the proliferation of EC cells, we generated the stable FBN1 knockdown cell line RL95-2-shFBN1 and stable double FBXO2/FBN1 double-knockdown cell line RL95-2-shFBXO2/shFBN1 (**Figure 5A**). We found that silencing of FBXO2 in RL95-2 cells significantly inhibited cell proliferation compared with RL95-2-NC. However, silencing of both FBXO2 and FBN1 in RL95-2 cells reversed this phenotype (**Figures 5B–D**). Flow cytometric assay revealed that knockdown of FBXO2 caused an increased apoptotic proportion of cells compared with the RL95-2-NC group and knockdown of both FBXO2 and FBN1 decreased the apoptotic proportion compared with the RL95-2-shFBXO2 group (**Figure 5E**). Flow cytometric assay also showed that knockdown of FBXO2 caused an increase in the percentage of cells in the G1 phase and a decrease in cells in the S phase compared with the RL95-2-NC group. Knockdown of both FBXO2 and FBN1 revealed a decrease in the percentage of cells in the G1 phase and a concomitant increase in cells in the G2 phase compared with the RL95-2-shFBXO2 group (**Figure 5F**). All these results demonstrated that FBN1 was required for FBXO2 to exert functional impacts on the proliferation of EC.

FBXO2 Regulates the Cell Cycle and the Autophagy Signaling Pathway

To explore the mechanisms underlying the regulatory actions of FBXO2 on EC cells, we performed RNA-seq analysis on RL95-2-NC and RL95-2-shFBXO2 cells and identified a total of 105 genes differentially expressed in RL95-2-shFBXO2 groups (with a padj < 0.05 and $|\text{Log2Fold change}| > 0$) (**Figure 6A**). By

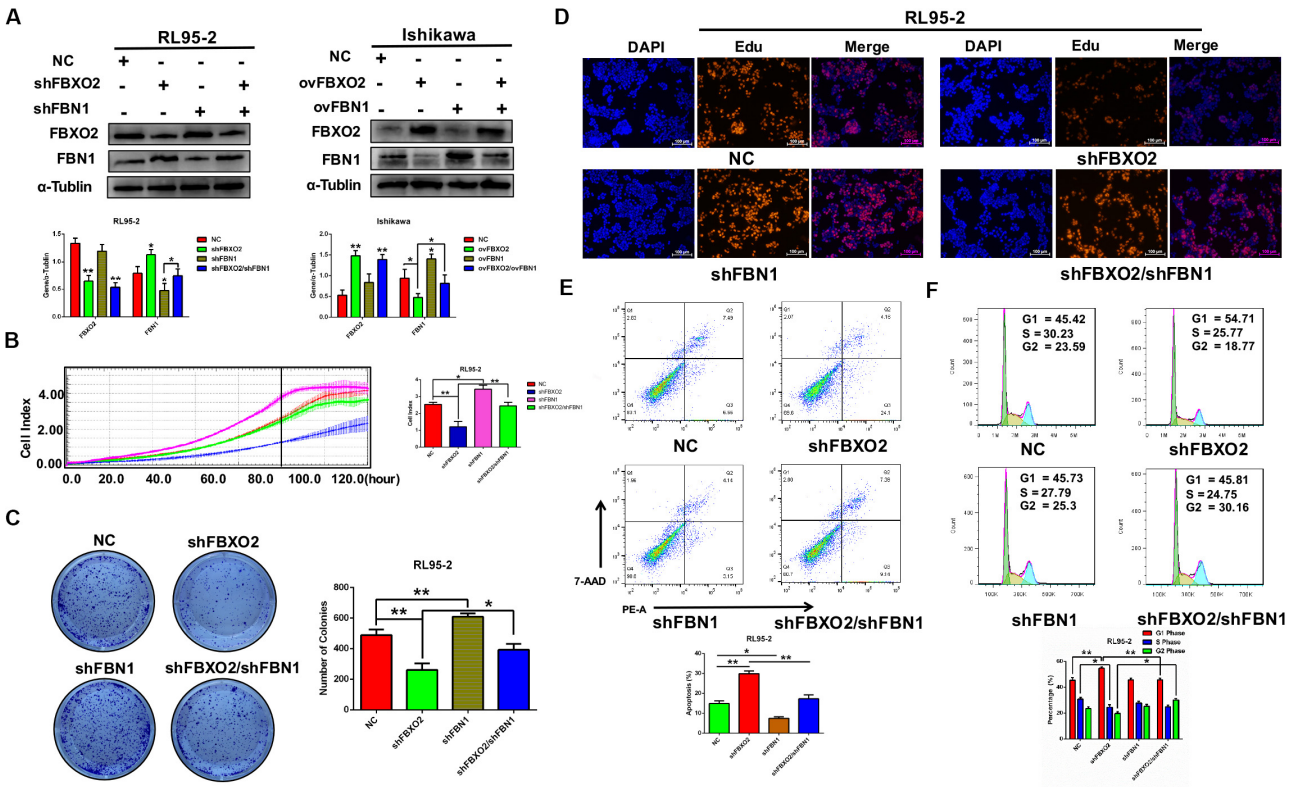


FIGURE 5 | FBXO2 regulates EC proliferation through FBN1. **(A)** FBXO2 and FBN1 protein levels in RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 stably transfected cell lines (left) and in Ishikawa-ovFBN1 and Ishikawa-ovFBXO2/ovFBN1 transiently transfected cell lines (right). *$P < 0.05$, **$P < 0.01$, one-way ANOVA. **(B,C)** Cell proliferation was measured by RTCA **(B)**, and colony formation assay **(C)** in RL95-2-shNC, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 stable cell lines. Data are presented as mean ± SD, *$P < 0.05$, **$P < 0.01$, one-way ANOVA. **(D)** EdU staining intensity and proportion in RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, RL95-2-shFBXO2/shFBN1 stable cell lines. **(E,F)** Apoptotic index **(E)** and cell cycle analysis **(F)** as measured by flow cytometric assay in RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 cell lines. Data are presented as mean ± SD, *$P < 0.05$, **$P < 0.01$, one-way ANOVA.

GO and KEGG pathway enrichment analysis, these genes were principally categorized into regulation of G1/S, cell cycle, focal adhesion, PI3k-Akt pathway, Wnt pathway, and p53 pathway (**Figures 6B,C**). By GSEA we categorized these genes into autophagy, and cell cycle among others (**Figures 6D–G**).

We showed that knockdown of FBXO2 caused a significant down-regulation of CDK4, CCND1, CCND2, and CCNA1 but not of MCM7, SMC1A, CHEK1, or CDC14B (**Figure 7A**). The western blotting results showed that knockdown of FBXO2 and overexpression of FBN1 decreased the protein expression of CDK4, Cyclin D1, Cyclin D2, and Cyclin A1 in RL95-2-shFBXO2 cells compared with RL95-2-NC cells. Knockdown of both FBXO2 and FBN1 significantly increased the expression of the aforementioned cell cycle proteins in RL95-2-shFBXO2/shFBN1 cells compared with RL95-2-shFBXO2 cells (**Figure 7B**). Furthermore, a significant upregulation of ATG4A and ATG4D (proteins involved in autophagic cell death) was observed in the shFBXO2 group compared with the RL95-2-NC group (**Figure 7C**). We also found that inhibition of cyclin-dependent kinases by palbociclib (PD0332991, which specifically inhibits cyclin-dependent kinases 4 and 6 [CDK4/6]) (Liu et al., 2018) reversed the effects of FBXO2 on proliferation of Ishikawa cells (**Figure 7D**). We also showed that the inhibition of 3-methyladenine (3-MA, an autophagy specific inhibitor) (Luan et al., 2019) reversed the effects of FBXO2 on proliferation of RL95-2 cells (**Figure 7E**). Hence, from our results, we proposed that FBXO2/FBN1 was important in promoting EC proliferation via the cell cycle and the autophagy signaling pathway.

FBXO2 Silencing Inhibits EC Carcinogenicity Through FBN1 *in vivo*

To support the *in vitro* findings, we evaluated the function of FBXO2 in EC using a nude mouse model. Mice were injected subcutaneously with RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, or RL95-2shFBXO2/shFBN1 stable cell lines at a concentration of 5×10^6 cells in 0.2 ml of medium. Compared with the mice injected with RL95-2-NC cells, those injected with RL95-2-shFBXO2 cells displayed an attenuated rate of tumor growth while knockdown of both FBXO2 and FBN1 rescued this phenotype (**Figures 7F,H,I**). In addition, weight of tumors derived from the RL95-2-shFBXO2 group was significantly reduced relative to that derived from the RL95-2-NC

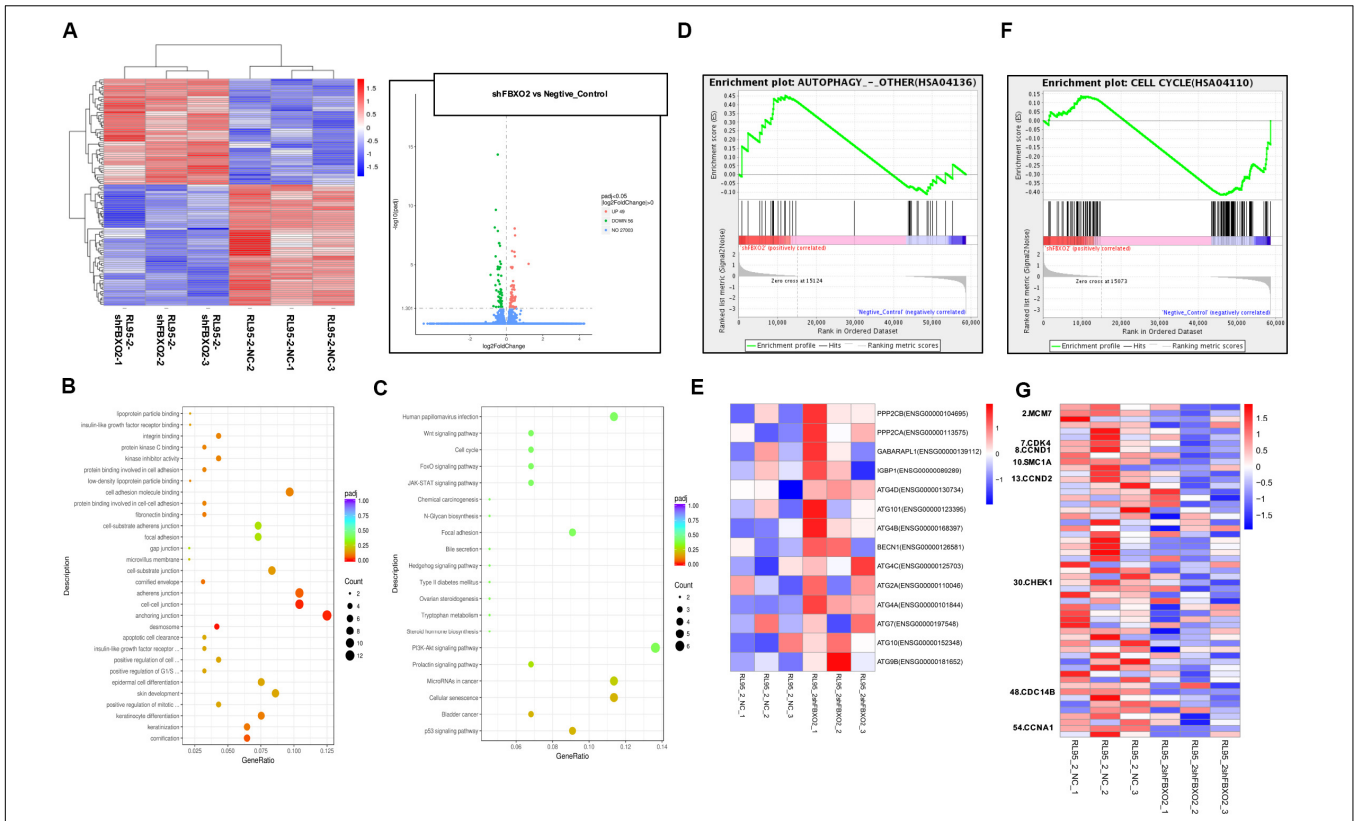


FIGURE 6 | FBXO2-FBN1 promote EC cells proliferation by regulating the cell cycle and autophagy signaling pathways. **(A)** Heatmap and volcano plot of differential genes enrichment based on the RNA sequencing of the RL95-2-NC and RL95-2-shFBXO2 groups. **(B)** Go functional enrichment analysis. **(C)** KEGG pathway enrichment analysis. **(D,E)** GSEA (autophagy signaling pathway) enrichment analysis and the corresponding heatmap. **(F,G)** GSEA (the cell cycle) enrichment analysis and the corresponding heatmap.

group. However, tumor weights were much higher with double-knockdown of FBXO2 and FBN1 compared with the RL95-2-shFBXO2 group (**Figure 7I**). When we used IHC to examine FBXO2, FBN1, and Ki67 expression in tumors, we confirmed that FBXO2 expression was diminished in the RL95-2-shFBXO2 group (**Figure 7J**, column 2) and that FBN1 was reduced in the RL95-2-shFBN1 group (**Figure 7J**, column 3). Knockdown of FBXO2 expression resulted in increased expression of FBN1 and decreased expression of Ki67 (**Figure 7J**, column 2), and knockdown of FBN1 or double-knockdown of FBXO2 and FBN1 resulted in decreased ki67 expression (**Figure 7J**, column 3 and **4**). This indicated that the absence of FBN1 blocked the proliferative role of FBXO2. FBXO2 was inversely correlated with FBN1 expression and positively correlated with Ki67 expression (**Figures 7J–L**). Collectively, these data indicate that FBXO2 functions as a tumor oncogene, and that it is essential for EC cell growth.

DISCUSSION

Using multiple complementary approaches, we identified that FBXO2, precisely the C-terminal FBA domain, directly binds to FBN1, leading to its degradation by polyubiquitination, promoting proliferation of EC by inactivating the cell cycle and inhibiting the autophagy signaling pathways.

F-box proteins regulate substrates in diverse biologic processed that control essential aspects of cellular life, including cell growth, development and differentiation, and cell survival and death. And dysregulation of F-box protein-mediated ubiquitylation has been demonstrated in various types of cancers (Frescas and Pagano, 2008; Crusio et al., 2010; Inuzuka et al., 2011). FBXO2, a member of the human F-box family, is a cytoplasmic protein and a neuron-enriched ubiquitin ligase substrate adaptor protein (Nelson et al., 2006) that binds the signature *N*-linked high-mannose glycan moiety of glycoproteins, and mediates the ubiquitination of ER glycoproteins in the ER-associated degradation system (Schroder and Kaufman, 2005). Furthermore, the FBXO2 variant rs99614 C allele was found to decrease the risk for Parkinson's disease in mainland Han Chinese (Yuan et al., 2017). FBXO2, along with the presence of an F box domain was also highly enriched in the nervous system, and involved in cell cycle regulation by directly interacting with Skp1p (Erhardt et al., 1998). In addition, the FBA domain of FBXO2 was equally essential for its activity of glycoprotein-recognizing activity. Liu et al. (2017) proved that FBXO2, along with the Fox-associated domain

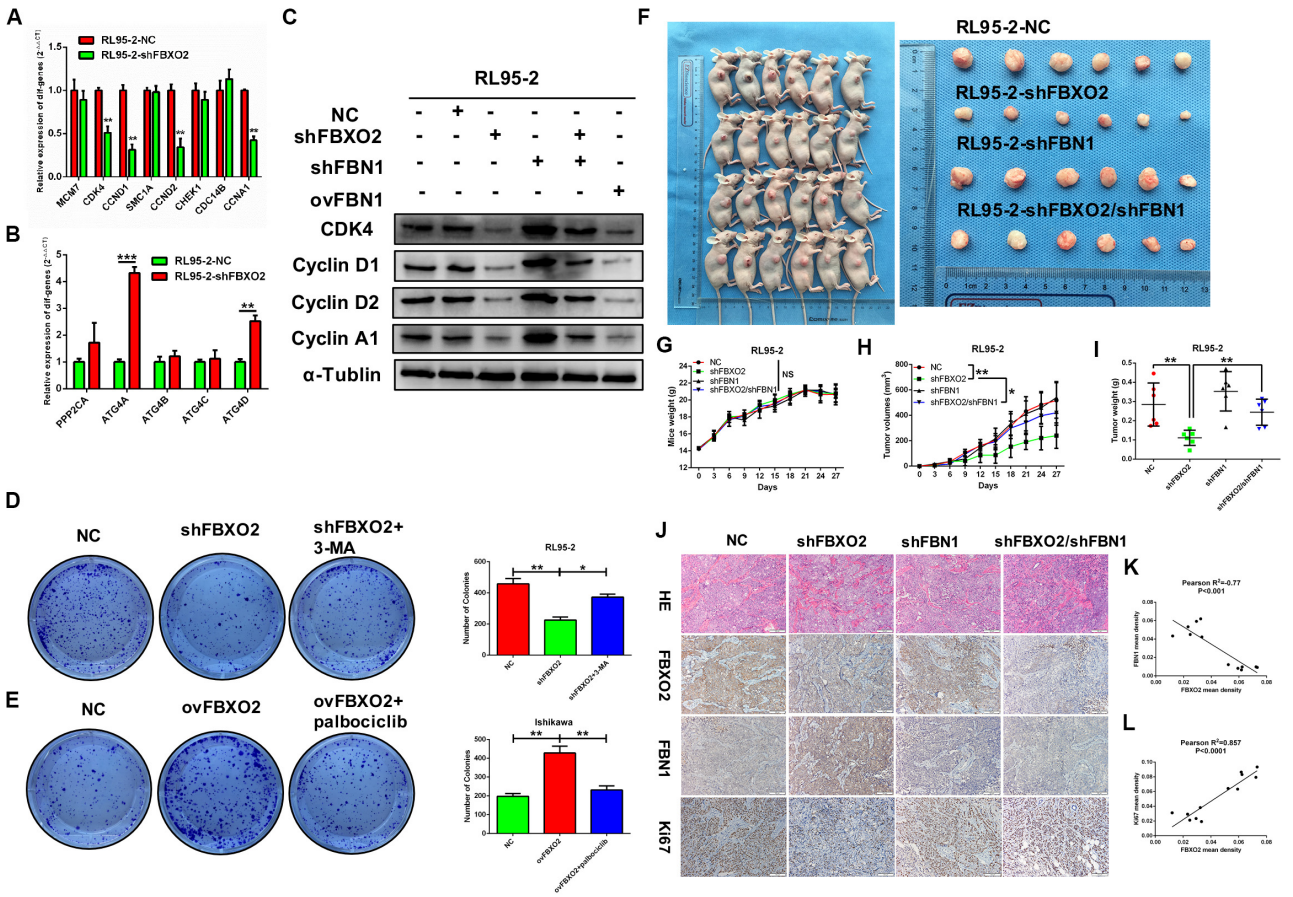


FIGURE 7 | (A–L) Verification of differential genes associated with the cell cycle and autophagy signaling pathways. **(A)** mRNA levels of MCM7, CDK4, CCND1, SMAC1A, CCND2, CHEK1, CDC14B, and CCNA1 in RL95-2-NC and RL95-2-shFBXO2 groups. CDK4, CCND1, CCND2, and CCNA1 were significantly down-regulated in the RL95-2-shFBXO2 group compared with RL95-2-NC group. Data are presented as mean ± SD, $^{**}P < 0.01$, *t*-test. **(B)** mRNA levels of PPP2CA, ATG4A, ATG4B, ATG4C, and ATG4D in the RL95-2-NC and RL95-2-shFBXO2 groups. ATG4A and ATG4D were significantly up-regulated in the RL95-2-shFBXO2 group compared with the RL95-2-NC group. Data are presented as mean ± SD, $^{**}P < 0.01$, $^{***}P < 0.001$, *t*-test. **(C)** Protein levels of CDK4, CyclinD1, CyclinD2, and CyclinA1 in RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, RL95-2-shFBXO2/shFBN1, and RL95-2-ovFBN1 cells. Knockdown of FBXO2 and overexpression of FBN1 revealed the same tendency in decreasing protein levels of CDK4, CyclinD1, CyclinD2, and CyclinA1 compared with the RL95-2-NC group. Knockdown of both FBXO2 and FBN1 showed increased protein levels of CDK4, CyclinD1, CyclinD2, and CyclinA1 compared with the RL95-2-shFBXO2 group. α-Tublin was used as a loading control. **(D)** The autophagy-specific inhibitor 3-MA (2 mM) reversed the effects of FBXO2 on the RL95-2 cells proliferation. Data are presented as mean ± SD, $^{*}P < 0.05$, $^{**}P < 0.01$, one-way ANOVA. **(E)** The CDK4/6 inhibitor palbociclib (2 μM) reversed the effects of FBXO2 on the Ishikawa cells proliferation. Data are presented as mean ± SD, $^{**}P < 0.01$, one-way ANOVA. **(F–L)**, FBXO2 silencing inhibited EC carcinogenicity *in vivo* via FBN1. **(F)** RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 cells were injected into BALB/c nude mice subcutaneously (0.2 ml, 5×10^6 cells) and harvested at day 27. **(G)** Mice weight were monitored every 3 days and recorded ($P > 0.05$). **(H,I)** Tumor weight was recorded in a time-dependent manner **(H)** and at harvest day **(I)** among the RL95-2-NC, RL95-2-shFBXO2, RL95-2-shFBN1, and RL95-2-shFBXO2/shFBN1 groups. Data are presented as mean ± SD, $^{*}P < 0.05$, $^{**}P < 0.01$, one-way ANOVA. **(J)** IHC staining for FBXO2, FBN1, and Ki67 in histologic sections of transplanted tumors. **(K)** Inverse correlation between FBXO2 and FBN1 staining, Pearson $r^2 = -0.77$, $P < 0.001$. **(L)** Positive correlation between FBXO2 and Ki67 staining, Pearson $r^2 = 0.857$, $P < 0.0001$.

(FBA) activity, interacted with the insulin receptor to enhance its ubiquitination-mediated protein degradation. FBXO2 also promoted the degradation of Epstein-Barr virus glycoprotein though its sugar-binding domain, and decreased the entry of the virus (Zhang et al., 2018). Investigators have in recent years studies the role of FBXO2 in various cancers. For example, FBXO2 was reported to regulate the EMT signaling pathway in gastric cancer (Sun et al., 2018) and high expression levels if FBXO2 correlated with colorectal cancer metastasis, such that it is now used as a reliable predictor of poor prognosis in colorectal cancer patients (Wei et al., 2018).

In this study, we found that FBXO2 was significantly up-regulated in EC compared with normal endometrium based on RNA-seq. We also determined the clinical relevance of FBXO2 in EC and demonstrated that FBXO2 levels were positively associated with tumor stage, grade, histologic type, and poor survival. Furthermore, FBXO2 showed high expression in EC samples compared with para-carcinoma tissues, and was positively associated with muscular infiltration and Ki67 staining (**Figure 1**). We also demonstrated that knockdown of FBXO2 inhibited EC cells proliferation, and that overexpression of FBXO2 promoted EC cells proliferation (**Figure 2**). We

used liquid chromatograph-tandem mass spectrometry and found that FBN1 was one of the substrates regulated by FBXO2 ubiquitination. We also demonstrated that FBN1 directly interacted with FBXO2 and showed ubiquitination-mediated degradation by FBXO2 (**Figures 3**, **4**). FBN1 is reported to be located on chromosome 15 and encodes a type of large extracellular matrix glycoprotein named fibrillin 1 (Biery et al., 1999). FBN1 can form threadlike filaments, and microfibrils, providing structural support for tissues, and forms elastic fibers in skin and blood vessels (Isogai et al., 2003; Massam-Wu et al., 2010). Prior studies have revealed that the FBN1 gene promoter was hypermethylated in colorectal cancer and endothelial tumor cells (Hellebrekers et al., 2007; Lind et al., 2011). Lack of the FBN1 protein, then, may play a potential role in tumors and the FBN1 gene may be considered an essential tumor-suppressor gene. In addition, plasma FBN1 has proven to be one promising biomarker in aiding the diagnosis of spontaneous coronary artery dissection (Zhu et al., 2018; Hui et al., 2020). Our next approach showed that FBN1 was required for FBXO2 to exert functional impacts on EC proliferation (**Figure 5**). The physiologic role of FBXO2 was further revealed with nude mice experiments (**Figure 7**), which suggested that inhibiting the expression or activity of FBXO2 represents a potential therapeutic role for the precise treatment of EC.

Our data also indicated that aberrant expression of FBXO2 was attributed to the cell cycle and autophagy signaling pathway (**Figures 6**, **7**). The cell cycle is the sequence of events by which a cell duplicates its genome, and grows, and divides. Cyclin-dependent kinases and other kinases comprise the key cell cycle transitions. Dysregulation of the cell cycle is a hallmark of cancer that leads to aberrant cellular proliferation (Poon, 2016; Bonelli et al., 2019). In our study, knockdown of FBXO2 or overexpression of FBN1 led to significant down-regulation of CDK1, CyclinD1, CyclinD2, and CyclinA1. Furthermore, knockdown of both FBXO2 and FBN1 reversed levels of the aforementioned proteins. Importantly, the CDK4/6 inhibitor, Palbociclib, reversed the effects of FBXO2 on proliferation in Ishikawa cells (**Figure 7**). Pharmacologic inhibitors of CDK4/6 have recently shown the promising activity in patients with cancers. Based on our data, we suggest that those patients showing a high expression of FBXO2 are potential candidates for CDK4/6 inhibitor treatment. Autophagy, a conserved self-catabolic process, enables the cells to remove and recycle the cytoplasmic contents, such as toxic molecules and invading microorganisms, and is comprised of five stages: initiation, nucleation, maturation, fusion with the lysosome, and degradation (Shintani and Klionsky, 2004; Saha et al., 2018). This process is strictly coordinated by autophagic core proteins encoded by more than 30 AuTophaGy-related (ATG) genes. ATG4 (consisting of four homologs: ATG4A, ATG4B, ATG4C, and ATG4D) plays an essential role in the regulation of the Atg8/microtubule-associated protein 1A/1B-light chain 3 (LC3) lipid conjugation system (Fu et al., 2019). In the present study, knockdown of FBXO2 up-regulated ATG4A and ATG4D, suggesting that silencing FBXO2 may activate ATG cell death. In addition, 3-MA (one autophagy specific inhibitor) reversed the effects of FBXO2 on the proliferation of RL95-2 cells (**Figure 7**).

In all, our results have identified FBXO2 as a functional ubiquitin enzyme for FBN1 and provided valuable insights into a potential role of FBXO2 in EC treatment. FBXO2 may be serve as a potential therapeutic target for EC precise treatment.

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 studies involving human participants were reviewed and approved by this present study was approved by the Ethics Committee of Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine and Obstetrics and Gynecology Hospital of Fudan University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by all animals were handled according to the Guide for the Care and Use of Laboratory Animals' and the Principles for the Utilization and Care of Vertebrate Animals.

AUTHOR CONTRIBUTIONS

WF: design of this study. XC, FJ, YW, and JZ: acquisition of data. XC, FJ, QC, and XW: data analysis. XC, FJ, and JS: drafting the article. WF and NJ: grammar and language modification. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Shanghai Municipal Education Commission-Gaofeng Clinical Medicine Grant (Grant No. 20172003); the National Natural Science Foundation of China (Grant No. 81572836); and the Science and Technology Commission of Shanghai Municipality (Grant No. 16411953500).

ACKNOWLEDGMENTS

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.00843/full#supplementary-material

REFERENCES

Biery, N. J., Eldadah, Z. A., Moore, C. S., Stetten, G., Spencer, F., and Dietz, H. C. (1999). Revised genomic organization of FBN1 and significance for regulated gene expression. *Genomics* 56, 70–77. doi: 10.1006/geno.1998.5697

Bonelli, M., La Monica, S., Fumarola, C., and Alfieri, R. (2019). Multiple effects of CDK4/6 inhibition in cancer: from cell cycle arrest to immunomodulation. *Biochem. Pharmacol.* 170:113676. doi: 10.1016/j.bcp.2019.113676

Cenciarelli, C., Chiaur, D. S., Guardavaccaro, D., Parks, W., Vidal, M., and Pagano, M. (1999). Identification of a family of human F-box proteins. *Curr. Biol.* 9, 1177–1179. doi: 10.1016/S0960-9822(00)80020-2

Che, X., Yan, H., Sun, H., Dongol, S., Wang, Y., Lv, Q., et al. (2016). Grifolin induces autophagic cell death by inhibiting the Akt/mTOR/S6K pathway in human ovarian cancer cells. *Oncol. Rep.* 36, 1041–1047. doi: 10.3892/or.2016.4840

Crusio, K. M., King, B., Reavie, L. B., and Aifantis, I. (2010). The ubiquitous nature of cancer: the role of the SCF(Fbw7) complex in development and transformation. *Oncogene* 29, 4865–4873. doi: 10.1038/onc.2010.222

Cuevas, I. C., Sahoo, S. S., Kumar, A., Zhang, H., Westcott, J., Aguilar, M., et al. (2019). Fbxw7 is a driver of uterine carcinosarcoma by promoting epithelial-mesenchymal transition. *Proc. Natl. Acad. Sci. U.S.A.* 116, 25880–25890. doi: 10.1073/pnas.1911310116

Erhardt, J. A., Hynicka, W., DiBenedetto, A., Shen, N., Stone, N., Paulson, H., et al. (1998). A novel F box protein, NFB42, is highly enriched in neurons and induces growth arrest. *J. Biol. Chem.* 273, 35222–35227. doi: 10.1074/jbc.273.52.35222

Frescas, D., and Pagano, M. (2008). Deregulated proteolysis by the F-box proteins SKP2 and beta-TrCP: tipping the scales of cancer. *Nat. Rev. Cancer* 8, 438–449. doi: 10.1038/nrc2396

Fu, Y., Huang, Z., Hong, L., Lu, J. H., Feng, D., Yin, X. M., et al. (2019). Targeting ATG4 in cancer therapy. *Cancers (Basel)* 11, 649. doi: 10.3390/cancers11050649

Glenn, K. A., Nelson, R. F., Wen, H. M., Mallinger, A. J., and Paulson, H. L. (2008). Diversity in tissue expression, substrate binding, and SCF complex formation for a lectin family of ubiquitin ligases. *J. Biol. Chem.* 283, 12717–12729. doi: 10.1074/jbc.M709508200

Hellebrekers, D. M., Melotte, V., Vire, E., Langenkamp, E., Molema, G., Fuks, F., et al. (2007). Identification of epigenetically silenced genes in tumor endothelial cells. *Cancer Res.* 67, 4138–4148. doi: 10.1158/0008-5472.CAN-06-3032

Hui, P., Bai, Y., Su, X., Quan, N., Qiao, B., Zheng, Y., et al. (2020). The value of plasma fibrillin-1 level in patients with spontaneous coronary artery dissection. *Int. J. Cardiol.* 302, 150–156. doi: 10.1016/j.ijcard.2019.12.015

Ilyin, G. P., Serandour, A. L., Pigeon, C., Rialland, M., Glaise, D., Guguen-Guillouzo, C., et al. (2002). A new subfamily of structurally related human F-box proteins. *Gene* 296, 11–20. doi: 10.1016/s0378-1119(02)00867-3

Inuzuka, H., Shaik, S., Onoyama, I., Gao, D., Tseng, A., Maser, R. S., et al. (2011). SCF(FBW7) regulates cellular apoptosis by targeting MCL1 for ubiquitylation and destruction. *Nature* 471, 104–109. doi: 10.1038/nature09732

Isogai, Z., Ono, R. N., Ushiro, S., Keene, D. R., Chen, Y., Mazzieri, R., et al. (2003). Latent transforming growth factor beta-binding protein 1 interacts with fibrillin and is a microfibril-associated protein. *J. Biol. Chem.* 278, 2750–2757. doi: 10.1074/jbc.M209256200

Kamata, Y., Watanabe, J., Nishimura, Y., Arai, T., Kawaguchi, M., Hattori, M., et al. (2005). High expression of skp2 correlates with poor prognosis in endometrial endometrioid adenocarcinoma. *J. Cancer Res. Clin. Oncol.* 131, 591–596. doi: 10.1007/s00432-005-0671-2

Lind, G. E., Danielsen, S. A., Ahlquist, T., Merok, M. A., Andresen, K., Skotheim, R. I., et al. (2011). Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. *Mol. Cancer* 10:85. doi: 10.1186/1476-4598-10-85

Liu, B., Lu, H., Li, D., Xiong, X., Gao, L., Wu, Z., et al. (2017). Aberrant expression of FBXO2 disrupts glucose homeostasis through Ubiquitin-Mediated degradation of insulin receptor in obese mice. *Diabetes* 66, 689–698. doi: 10.2337/db16-1104

Liu, J., Peng, Y., Zhang, J., Long, J., Liu, J., Wei, W., et al. (2020). Targeting SCF e3 ligases for cancer therapies. *Adv. Exp. Med. Biol.* 1217, 123–146. doi: 10.1007/978-981-15-1025-0_9

Liu, M., Liu, H., and Chen, J. (2018). Mechanisms of the CDK4/6 inhibitor palbociclib (PD 0332991) and its future application in cancer treatment (Review). *Oncol. Rep.* 39, 901–911. doi: 10.3892/or.2018.6221

Luan, W., Pang, Y., Li, R., Wei, X., Jiao, X., Shi, J., et al. (2019). Akt/mTOR-Mediated autophagy confers resistance to BET inhibitor JQ1 in ovarian cancer. *Oncol. Targets Ther.* 12, 8063–8074. doi: 10.2147/OTT.S220267

Massam-Wu, T., Chiu, M., Choudhury, R., Chaudhry, S. S., Baldwin, A. K., McGovern, A., et al. (2010). Assembly of fibrillin microfibrils governs extracellular deposition of latent TGF beta. *J. Cell Sci.* 123, 3006–3018. doi: 10.1242/jcs.073437

Mizushima, T., Yoshida, Y., Kumanomidou, T., Hasegawa, Y., Suzuki, A., Yamane, T., et al. (2007). Structural basis for the selection of glycosylated substrates by SCF(Fbs1) ubiquitin ligase. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5777–5781. doi: 10.1073/pnas.0610312104

Morice, P., Leary, A., Creutzberg, C., Abu-Rustum, N., and Darai, E. (2016). Endometrial cancer. *Lancet* 387, 1094–1108. doi: 10.1016/S0140-6736(15)00130-0

Nelson, R. F., Glenn, K. A., Miller, V. M., Wen, H., and Paulson, H. L. (2006). A novel route for F-box protein-mediated ubiquitination links CHIP to glycoprotein quality control. *J. Biol. Chem.* 281, 20242–20251. doi: 10.1074/jbc.M602423200

Poon, R. Y. (2016). Cell cycle control: a system of interlinking oscillators. *Methods Mol. Biol.* 1342, 3–19. doi: 10.1007/978-1-4939-2957-3_1

Saha, S., Panigrahi, D. P., Patil, S., and Bhutia, S. K. (2018). Autophagy in health and disease: a comprehensive review. *Biomed. Pharmacother.* 104, 485–495. doi: 10.1016/j.biopha.2018.05.007

Schroder, M., and Kaufman, R. J. (2005). ER stress and the unfolded protein response. *Mutat. Res.* 569, 29–63. doi: 10.1016/j.mrfmmm.2004.06.056

Shintani, T., and Klionsky, D. J. (2004). Autophagy in health and disease: a double-edged sword. *Science* 306, 990–995. doi: 10.1126/science.1099993

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J. Clin.* 69, 7–34. doi: 10.3322/caac.21551

Skaar, J. R., Pagan, J. K., and Pagano, M. (2013). Mechanisms and function of substrate recruitment by F-box proteins. *Nat. Rev. Mol. Cell Biol.* 14, 369–381. doi: 10.1038/nrm3582

Sun, X., Wang, T., Guan, Z. R., Zhang, C., Chen, Y., Jin, J., et al. (2018). FBXO2, a novel marker for metastasis in human gastric cancer. *Biochem. Biophys. Res. Commun.* 495, 2158–2164. doi: 10.1016/j.bbrc.2017.12.097

Uddin, S., Bhat, A. A., Krishnankutty, R., Mir, F., Kulinski, M., and Mohammad, R. M. (2016). Involvement of F-BOX proteins in progression and development of human malignancies. *Semin. Cancer Biol.* 36, 18–32. doi: 10.1016/j.semcancer.2015.09.008

Vucic, D., Dixit, V. M., and Wertz, I. E. (2011). Ubiquitylation in apoptosis: a post-translational modification at the edge of life and death. *Nat. Rev. Mol. Cell Biol.* 12, 439–452. doi: 10.1038/nrm3143

Wei, X., Bu, J., Mo, X., Lv, B., Wang, X., and Hou, B. (2018). The prognostic significance of FBXO2 expression in colorectal cancer. *Int. J. Clin. Exp. Pathol.* 11, 5054–5062.

Yan, L., Lin, M., Pan, S., Assaraf, Y. G., Wang, Z. W., and Zhu, X. Q. (2019). Emerging roles of F-box proteins in cancer drug resistance. *Drug Resist. Updat.* 49:100673. doi: 10.1016/j.drup.2019.100673

Yoshida, Y., Chiba, T., Tokunaga, F., Kawasaki, H., Iwai, K., Suzuki, T., et al. (2002). E3 ubiquitin ligase that recognizes sugar chains. *Nature* 418, 438–442. doi: 10.1038/nature00890

Yuan, L., Song, Z., Deng, X., Yang, Z., Yang, Y., Guo, Y., et al. (2017). Genetic analysis of FBXO2, FBXO6, FBXO12, and FBXO41 variants in han chinese patients with sporadic parkinson's disease. *Neurosci. Bull.* 33, 510–514. doi: 10.1007/s12264-017-0122-5

Zhang, H. J., Tian, J., Qi, X. K., Xiang, T., He, G. P., Zhang, H., et al. (2018). Epstein-Barr virus activates F-box protein FBXO2 to limit viral infectivity by targeting glycoprotein B for degradation. *PLoS Pathog.* 14:e1007208. doi: 10.1371/journal.ppat.1007208

Zhi, X., Wang, Y., Zhou, X., Yu, J., Jian, R., Tang, S., et al. (2010). RNAi-mediated CD73 suppression induces apoptosis and cell-cycle arrest in human breast cancer cells. *Cancer Sci.* 101, 2561–2569. doi: 10.1111/j.1349-7006.2010.01733.x

Zhu, Z., Tang, W., Ge, L., Han, X., and Dong, Q. (2018). The value of plasma fibrillin-1 level in patients with spontaneous cerebral artery dissection. *Neurology* 90, e732–e737. doi: 10.1212/WNL.0000000000005027

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 © 2020 Che, Jian, Wang, Zhang, Shen, Cheng, Wang, Jia and Feng. 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 Oncology

ORIGINAL RESEARCH
published: 02 September 2020
doi: 10.3389/fonc.2020.01680

NEDD4 Plays Roles in the Maintenance of Breast Cancer Stem Cell Characteristics

*Seon-Ae Jeon, Dong Wook Kim, Da-Bin Lee and Je-Yoel Cho**

Department of Veterinary Biochemistry, BK21 Plus and Research Institute for Veterinary Science, School of Veterinary Medicine, Seoul National University, Seoul, South Korea

Triple-negative breast cancer (TNBC) is the most aggressive type with poor prognosis among the breast cancers and has a high population of cancer stem cells (CSCs), which are the main target to cure and inhibit TNBC. In this study, we examined the role of neural precursor cell expressed developmentally downregulated protein 4 (NEDD4) in the proliferation, migration, and CSC characteristics of MDA-MB-231, a TNBC cell line. Interestingly, the Kaplan–Meier plotter showed that the survival rate of patients with a higher expression level of NEDD4 was significantly shorter than those of patients with a lower expression only in relatively aggressive and higher stage (grade 3) breast cancer patients. The knockdown of NEDD4 drastically decreased the proliferation, migration, and mammosphere formation in MDA-MB-231 cells. A proteomic analysis revealed the alteration of CSC-related proteins; notably, Myc targets stem cell-like signatures in siNEDD4-treated MDA-MB-231. An immunoassay also showed that the expression and the activity of breast CSC markers are decreased in NEDD4-deleted MDA-MB-231. Taken together, these results indicate that NEDD4 is involved in the maintenance of populations and characteristics of breast CSCs.

Keywords: TNBC, BCSC, NEDD4, LC–MS/MS, TMT labeling

OPEN ACCESS

Edited by:
Harsha Gowda, The University of Queensland, Australia

Reviewed by:
Prasanna Ekambaram, University of Pittsburgh, United States
Antonietta Santoro, University of Salerno, Italy

****Correspondence:***
Je-Yoel Cho
jeycho@snu.ac.kr

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *06 April 2020*
Accepted: *29 July 2020*
Published: *02 September 2020*

Citation:
Jeon S-A, Kim DW, Lee D-B and Cho J-Y (2020) NEDD4 Plays Roles in the Maintenance of Breast Cancer Stem Cell Characteristics. Front. Oncol. 10:1680. doi: 10.3389/fonc.2020.01680

INTRODUCTION

Breast cancer (BC) is the most common and malignant cancer in women (1) and classified into three subtypes depending on certain molecular biomarkers (2): estrogen receptor (ER)/progesterone receptor (PR)-positive, HER2-positive, and triple-negative breast cancers (TNBCs). The ER/PR-positive breast cancers account for more than 60% of breast cancer patients. The HER2-positive breast cancers comprise 10–20% of breast cancer patients and have a worse diagnosis than ER/PR-positive breast cancers (3). However, these two types of breast cancers are responsive to ER/PR and HER2-targeted therapy, which blocks their downstream signal activation. TNBC, which lacks ER/PR and HER2 on the surface of the cancer cells, occupies 15–20% of all breast cancer patients and has a worse outcome compared to ER/PR-positive and HER2-positive subtypes (4). Recently, a report suggested that TNBC can be classified into six subtypes (immunomodulatory, mesenchymal, mesenchymal stem-like, luminal androgen receptor, and two basal-like subtypes) by gene expression profile (5). Among the six subtypes of TNBCs, the stem cell-like group has a completely distinguished expression pattern of molecular signatures compared to ER/PR- and HER2-positive breast cancers.

The specific subset of cells termed cancer stem cells (CSCs) in both hematologic malignancies and solid tumors has the propensity for tumor initiation and recurrence. The population of these specific cells is very small in tumors, and they have the unique capacity for self-renewal, differentiation, invasion, and drug resistance (6, 7). It has been well described that breast cancer stem cells (BCSCs) also have specific properties in common with normal mammary stem cells and progenitor cells (8). Although there have been several proposed BCSC markers, the major three markers are a cluster of differentiation 44-positive ($CD44^+$), CD24-negative ($CD24^-$), and aldehyde dehydrogenase-positive 1 ($ALDH1^+$) (9, 10). Interestingly, TNBCs have a relatively abundant BCSC population compared to other types of breast cancer, which makes these tumors more aggressive and drug resistant.

Neural precursor cell expressed developmentally downregulated protein 4 (NEDD4) is an E3 ubiquitin ligase which regulates target protein stability and cellular localization *via* proteosomal degradation (11). NEDD4 plays a critical role in diverse cellular functions in cancers, including tumor initiation, progression, migration, and resistance to anticancer therapies (12, 13). The expression of NEDD4 is frequently upregulated in several human cancers (13–15) and positively correlated with cell proliferation and survival *via* ubiquitination-mediated proteosomal degradation of tumor suppressors, such as PTEN (16, 17), and LATS (18, 19). In addition, NEDD4 exerts its oncogenic activity by stabilizing the mouse double minute 2 homolog that is also a ubiquitin ligase for p53, a tumor suppressor (20). Moreover, many other cancer-related signaling pathways are also regulated by NEDD4, such as pAKT, IGF1R, and NOTCH pathways (21–24). Given the role of NEDD4 in cancer, NEDD4 is considered to be a promising therapeutic target for the treatment of human malignancies (25). Although many signaling pathways regulated by NEDD4 in cancers have been discovered, the role of NEDD4 in CSCs remains elusive.

In this study, we analyzed the CSC-related proteome alteration by NEDD4 knockdown in CSC-abundant MDA-MB-231 cells by liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis. Then, we studied the effect of NEDD4 knockdown on the proliferation and the migration of MDA-MB-231 cells and the expression and the activity of the BCSC markers to reveal the role of NEDD4 in breast CSC maintenance.

MATERIALS AND METHODS

Cell Culture and siRNA Transfection

MCF7, SKBR3, and MDA-MB-231 were purchased from Korea Cell Line Bank (Seoul, South Korea) and MDA-MB-436 was purchased from the American Type Culture Collection (Manassas, VA, United States). MCF7, MDA-MB-231, and MDA-MB-436 were cultured in Dulbecco's modified Eagle medium with 10% fetal bovine serum (FBS), and SKBR3 was cultured in RPMI-1640 with 10% FBS in a humidified incubator at 37°C and 5% CO_2. The transfection experiments of siRNAs were performed using Lipofectamine 3000 (Invitrogen, *Carlsbad,* CA, United States). NEDD4 siRNA was purchased from Bioneer (Daejeon, South Korea). The following siRNA oligonucleotides were used in the transfection studies: siNEDD4 (target sequence, 5′-ATGGAGTTGATTAGATTACAA-3′), 5′-GGAGUUGAUUAGAUUACAATT-3′ (sense strand), and 5′-UUGUAAUCUAAUCAACUCCAT-3′ (antisense strand).

Quantitative PCR

Quantitative PCR was performed as previously reported (26). Briefly, total RNA was isolated from MDA-MB-231 using TRIzol reagent. Two micrograms of RNA was reverse-transcribed with an Omniscript reverse transcriptase kit for 1 h at 37°C. The differences in the mRNA expression of the target genes were determined by SYBR Green I fluorescence-based qRT-PCR. The mRNA level of each gene was normalized to *GAPDH*. The sequences of the primers are shown in **Supplementary Table 1**.

Protein Isolation and Western Blotting

Protein isolation and western blotting were done as previously reported (26). Each protein band was normalized against β-actin. Anti-NEDD4 antibody (ab14592, Abcam; Cambridge, Cambridgeshire, United Kingdom) was used at 10,000:1 dilution, while anti-ALDH1A1 (ab134188) and anti-CD44 antibodies (ab157107, Abcam; Cambridge, Cambridgeshire, United Kingdom) as well as anti-β-actin antibody (A2228, Sigma Aldrich; St. Louis, MO, United States) were used at 1,000:1 dilution. The membrane was developed using a chemiluminescence detection system (ATTO Corporation, Tokyo, Japan).

Filter-Aided Sample Preparation and Tandem Mass Tag Labeling

To prepare tryptic peptides for LC–MS/MS analysis, filter-aided sample preparation (FASP) was performed according to a previous report on siCONT- and siNEDD4-treated MDA-MB-231, respectively. Briefly, the cells were lysed by SDT lysis buffer [4% SDS, 100 mM Tris–HCl (pH 7.6), 0.1 M DTT]. Then, a total of 250 μg of proteins extracted by urea buffer (8 M urea in 0.1 M Tris–HCl, pH 8.5) was alkylated and digested by trypsin in FASP filters on the shaking incubator at 250 rpm overnight at 37°C. The digests were dried by centrifugation under vacuum. For a quantitative comparison, tandem mass tag (TMT) six-plex labeling was performed with 80 μg of tryptic peptides according to the manufacturer's instruction (Thermofisher Scientific, Waltham, MA, United States).

LC–MS/MS Analysis

Spectra raw data were acquired on Orbitrap Fusion Lumos (Thermofisher Scientific, Waltham, MA, United States) with EASY-nLC 1200 (Thermofisher Scientific, Waltham, MA, United States). An auto-sampler was used to load 10-μl aliquots of the peptide solutions into an EASY column—Acclaim PepMap^{TM} 100 of i.d. 75 μm, length 2 cm, and particle size of 3 μm (Thermofisher Scientific, Waltham, MA, United States). Then, the trapped peptides were separated on an EASY-Spray Column—C_{18} analytic-column of i.d. 75 μm, length 500 mm, and 2-μm particle size (100 Å from Thermofisher Scientific, Waltham, MA, United States). The mobile phases were composed of 100% water (buffer A) and 100% acetonitrile (buffer B), and

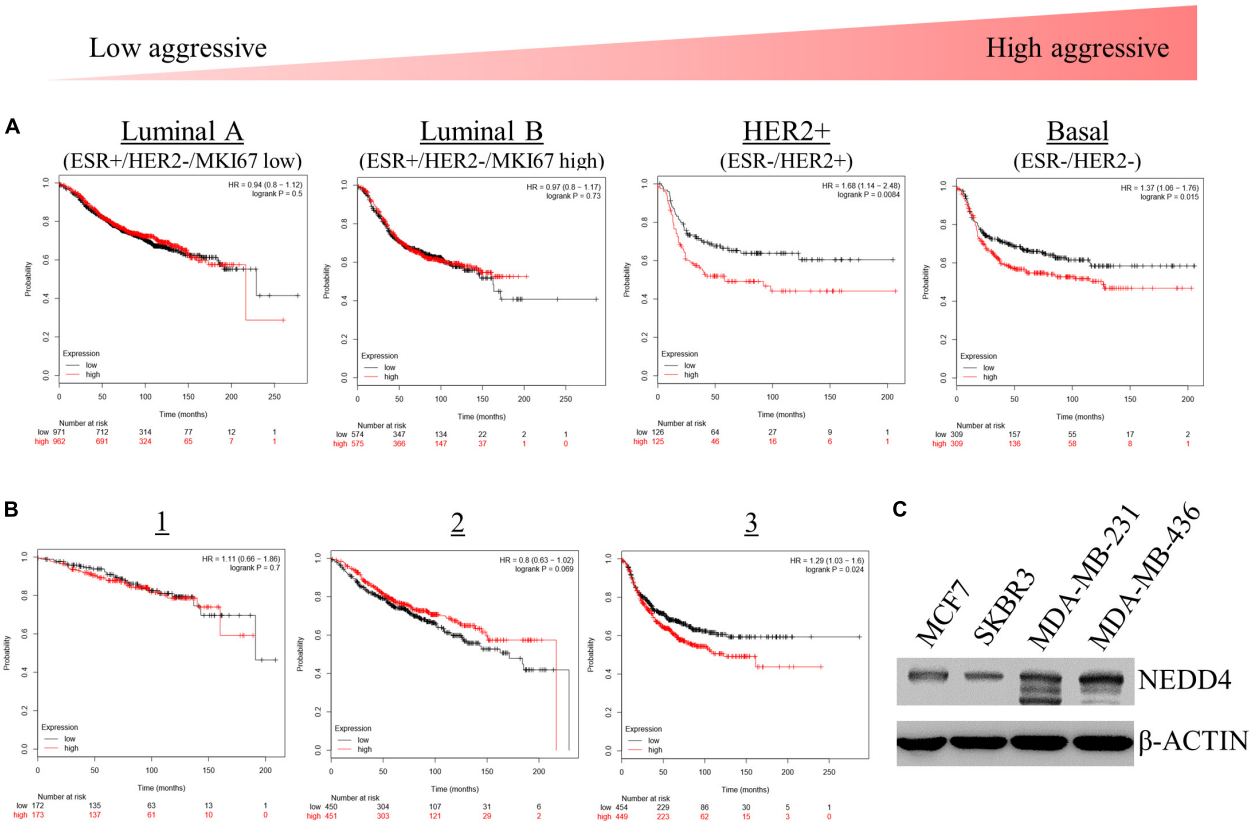


FIGURE 1 | The expression level of NEDD4 is negatively correlated with the survival rate in aggressive breast cancer patients. **(A,B)** The Kaplan–Meier survival curve showing the correlation between the level of NEDD4 expression and the aggressiveness and stage in breast cancer patients. **(C)** The expression of NEDD4 was confirmed in human breast cancer cell lines by western blot with an anti-NEDD4 antibody. β-actin was used as a loading control.

each contained 0.1% formic acid. The LC gradient was initiated with 5% buffer B, increased to 8% buffer B over 1 min, 10% buffer B over 16 min, 40% buffer B over 79 min, and then maintained at 80% buffer B for 9 min and 2% buffer B for an additional 15 min at a flow rate of 300 nl/min. During the chromatographic separation, the Orbitrap Fusion Lumos was operated in a data-dependent acquisition mode. Survey full scans were acquired on mass range 400–1,600 m/z, maximum injection time of 100 ms, and automatic gain control (AGC) target 2e5 ions with resolution of 120,000 and analyzed using the Orbitrap. The MS/MS precursors were selected from top *n* intense ions in 3 s between the survey scans, which were fragmented by 37.5% higher collisional dissociation collision energy. The MS/MS data were acquired on a maximum injection time of 54 ms and AGC 5e4 ions with a resolution of 30,000 and analyzed using the Orbitrap. The previously fragmented precursors were excluded for 30 s.

Protein Identification and Quantitative Data Analysis

Raw MS spectra were processed with MaxQuant software (version 1.5.8.3) (27) at default settings with unique peptide ≥ 2 and minimum number of amino acid ≥ 6. Identified peaks were searched against a database of human cancer from Uniprot[1]. The output files generated from Maxquant were subjected into Scaffold Q + software (version Scaffold_4.7.5, Proteome Software Inc., Portland, OR, United States) to TMT-labeled peptide and protein identifications. The peptide identifications were accepted if they could be established at greater than 5.0% probability to achieve a false discovery rate (FDR) of less than 1.0% by the Scaffold Local FDR algorithm. The protein identifications were accepted if they could be established at greater than 59.0% probability to achieve a FDR of less than 1.0% and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (28). The proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing a significant peptide evidence were grouped into clusters. Normalization was performed iteratively (across samples and spectra) on intensities, as described in the statistical analysis of relative labeled mass spectrometry data from complex samples using ANOVA (29). Medians were used for averaging. The spectra data were log-transformed, pruned of those matched to multiple proteins and

[1]https://www.uniprot.org/

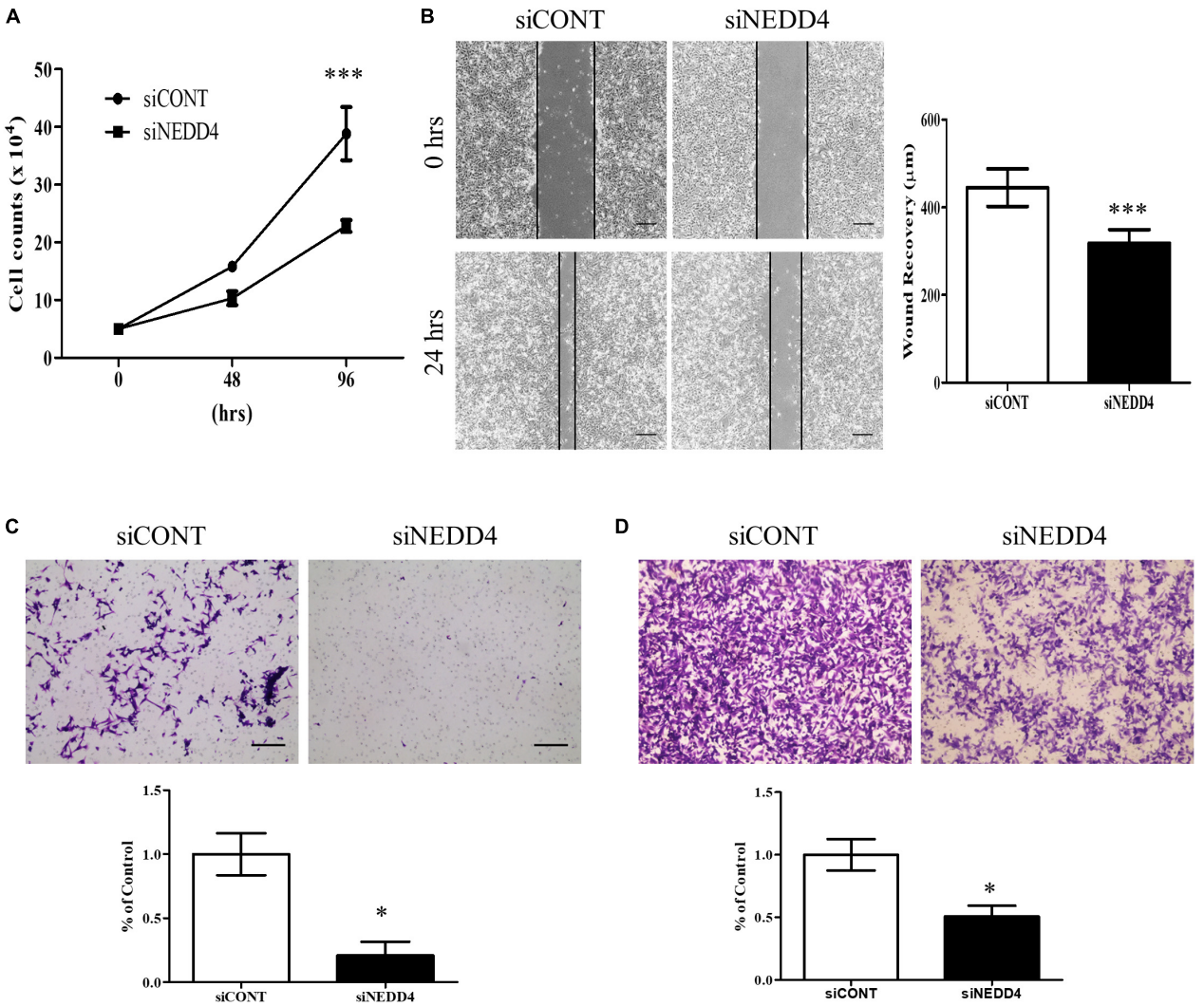


FIGURE 2 | Cell proliferation and migration are decreased in NEDD4-deleted MDA-MB-231. **(A)** The cell proliferation rate was confirmed by cell counting at 48 and 96 h after transfection of control siRNA (siCONT) and siNEDD4 into MDA-MB-231 ($n = 3$). **(B)** The wound healing assay was performed in siCONT- and siNEDD4-treated MDA-MB-231. At 24 h after wounding, the images were captured to measure the distance that was filled through cell migration and proliferation. The representative images are shown. The average value is shown in the graph ($n = 3$). Scale bar, 10 μm. **(C)** Cell migration assay was performed with Transwell cell culture chambers in siCONT- and siNEDD4-treated MDA-MB-231. Scale bar, 10 μm. **(D)** The invasiveness capacity was confirmed with matrigel-coated transwell chambers in siCONT- and siNEDD4-treated MDA-MB-231. The representative images are shown ($n = 3$) (*$P < 0.05$ and ***$P < 0.001$).

those missing a reference value, and weighted by an adaptive-intensity weighting algorithm.

Wound Healing Assay

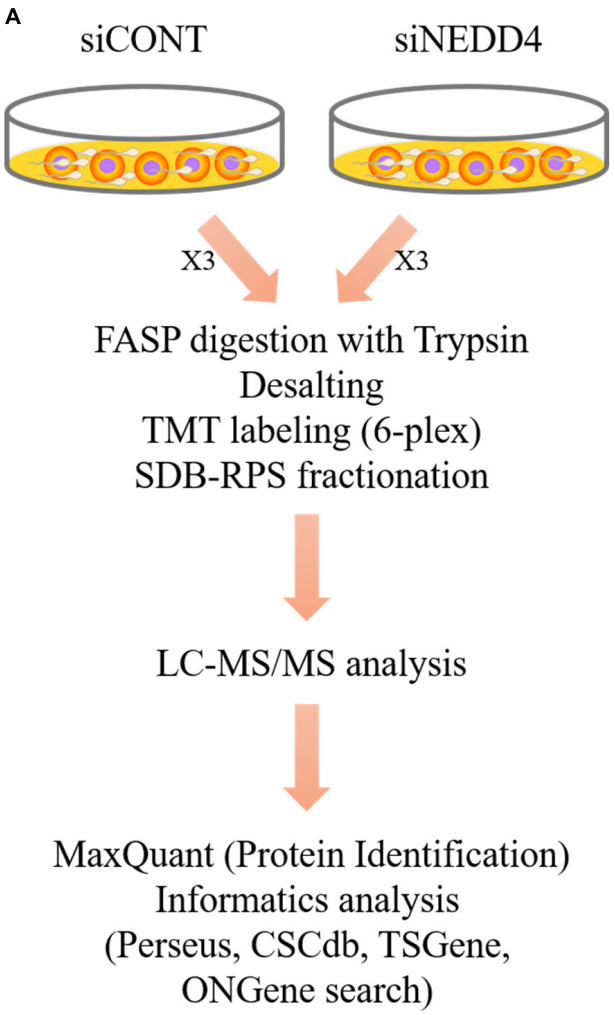
The cells treated with siCONT and siNEDD4 were, respectively, wounded with SPL Scarnd (SPL, Pocheon, Gyeonggido, South Korea), and the media were changed to freshly grown media. After 24 h, pictures were taken by microscopy (Revolve, ECHO, San Diego, CA, United States).

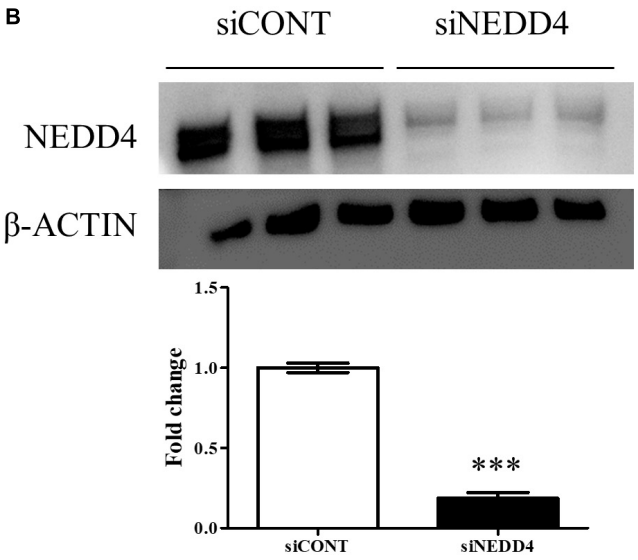
Migration Assay

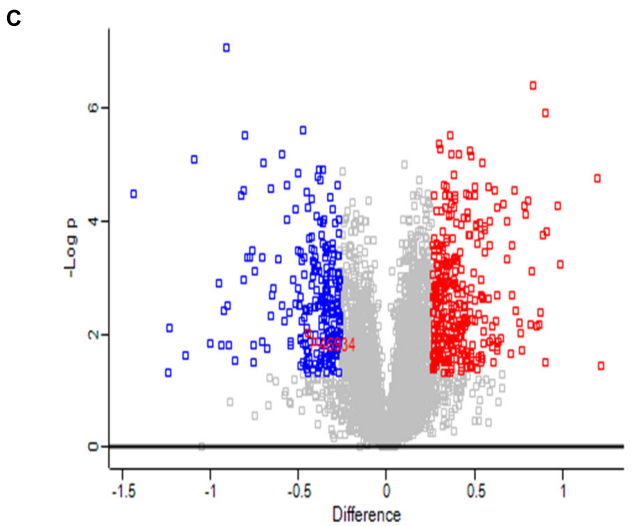
Cell migration ability was tested using 8.0-μm-pore polycarbonate-membrane-inserted transwell chambers (Corning, New York, NY, United States). The polycarbonate membrane of the upper chamber was coated with 0.1% gelatin, before the upper chambers with siCONT- and siNEDD4-treated cells (5×10^4 for each chamber). After 3 h, the cells through the membrane were fixed with 4% paraformaldehyde for 5 min and stained with 1% crystal violet in 2% ethanol. Pictures of the stained cells were taken by microscopy (Revolve, ECHO, San Diego, CA, United States).

Invasion Assay

Invasiveness ability was also tested using 8.0-μm-pore polycarbonate membrane-inserted transwell chambers (Corning, New York, NY, United States). The polycarbonate membrane of the upper chamber was coated with 0.3 mg/ml of matrigel, before the upper chambers with siCONT- and siNEDD4-treated







D

Gene Set enriched	SIZE	ES	NES	NOM p-val	FDR q-val
MYC_TARGETS_V2	42	-0.61583203	-1.2123097	0	0.2277616
E2F_TARGETS	126	-0.2878785	-1.2362523	0	0.23801604
MYC_TARGETS_V1	173	-0.29449773	-1.2636245	0	0.28299114
SPERMATOGENESIS	27	-0.5085487	-1.3686347	0	0.31849918
DNA_REPAIR	88	-0.37691373	-1.276257	0	0.35465607

FIGURE 3 | Identification of differentially expressed proteins (DEPs) by liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis in NEDD4-deleted MDA-MB-231. **(A)** Schematic diagram showing the overall workflow for proteomics analysis in siCONT- and siNEDD4-treated MDA-MB-231. The peptides obtained by trypsin digestion were labeled by tandem mass tag six-plex reagents. Then, identification and functional analysis of DEPs were performed by LC–MS/MS and bioinformatics tools. **(B)** The efficiency of NEDD4 knockdown by transfection of siNEDD4 into MDA-MB-231 was confirmed by western blot with an anti-NEDD4 antibody. β-actin was used as a loading control. **(C)** The volcano plot obtained from "Perseus" shows differentially expressed proteins in siNEDD4-transfected MDA-MB-231 compared to siCONT-transfected ones. Proteins in red or blue were considered as significantly up- or downregulated by using Perseus software. **(D)** Gene set enrichment analysis showing the gene sets enriched in downregulated proteins by NEDD4 knockdown (***$P < 0.001$).

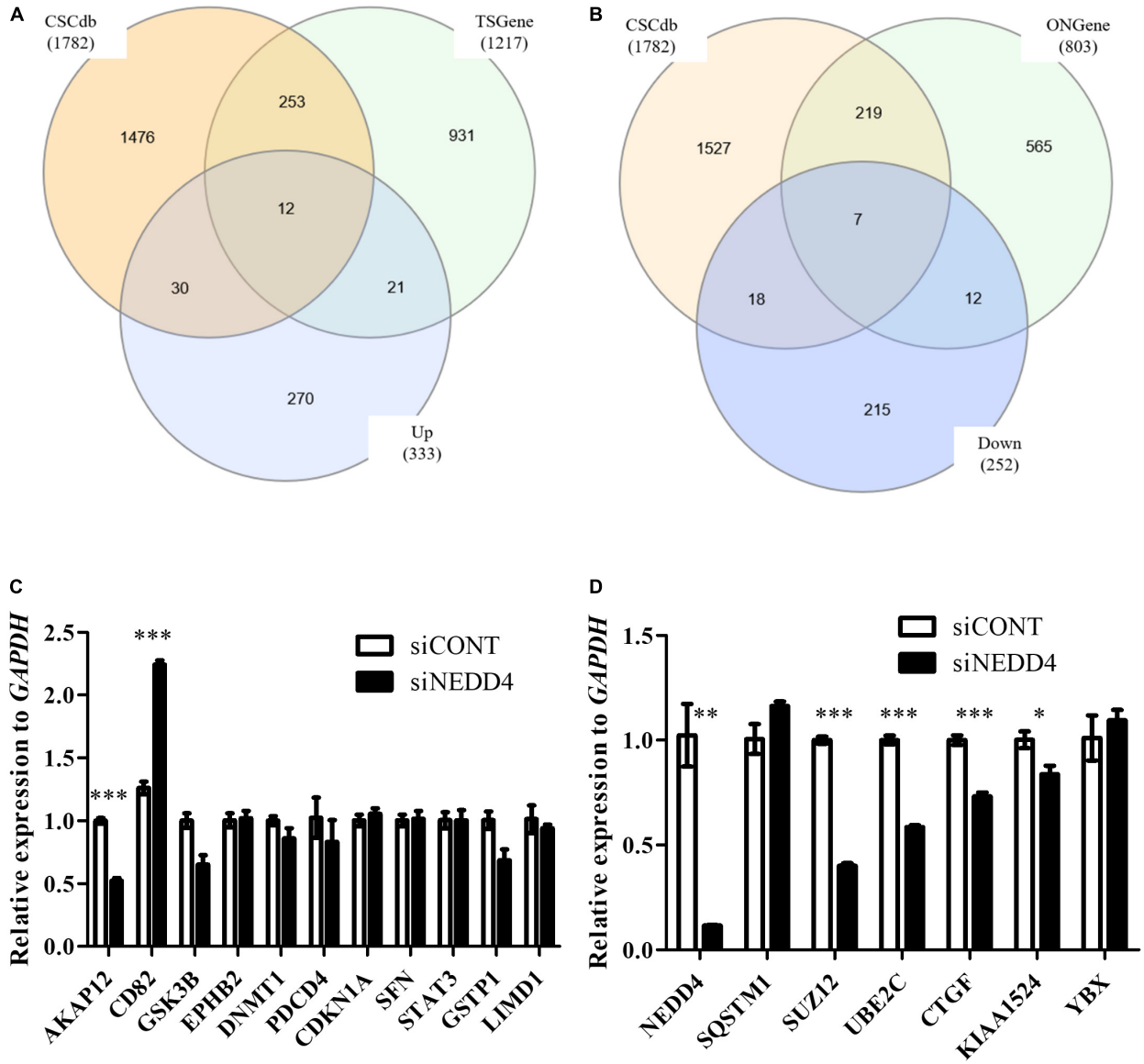


FIGURE 4 | A proteomic analysis reveals that the expression of cancer stem cell (CSC)-related proteins is altered by NEDD4 deletion. **(A)** Venn diagram showing the overlaps of upregulated proteins with CSC-related (CSCdb) and tumor suppressors (TSGene). **(B)** Venn diagram showing the overlaps of downregulated proteins with CSC-related (CSCdb) and oncogenes (ONGene). **(C)** The mRNA expression of CSC-related and increased proteins by NEDD4 knockdown was confirmed by qPCR in siCONT/NEDD4-treated MDA-MB-231, respectively. **(D)** The mRNA expression of CSC-related and decreased proteins by NEDD4 knockdown was confirmed by qPCR in siCONT/NEDD4-treated MDA-MB-231, respectively (*$P < 0.05$; **$P < 0.01$, and ***$P < 0.001$).

cells (5×10^4 for each chamber). After 24 h, cells through the membrane were fixed with 4% paraformaldehyde for 5 min and stained with 1% crystal violet in 2% ethanol. Pictures of the stained cells were taken by microscopy (Revolve, ECHO, San Diego, CA, United States).

ALDEFLUOR Assay

For the evaluation of ALDH1A1 activity, an ALDEFLUOR kit (STEMCELL Technologies, Vancouver, BC, Canada) was used according to the manufacturer's instructions. Briefly, siCONT/NEDD4-treated cells were cultured in growth media supplemented with 10 μl ALDEFLUOR reagent for 30 min, and then green fluorescence was detected by microscopy (Revolve, ECHO, San Diego, CA, United States).

Statistical Analysis

The data represent the means ± SEM. GraphPad Prism software (version 7.01, GraphPad software, Inc., CA, United States) was used for statistical analyses. The statistical significance of the results was assessed using Student's *t* test (two-tailed). For all experiments, significance was defined as *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

TABLE 1 | Cancer stem cell-related and upregulated proteins by NEDD4 knockdown in MDA-MB-231.

Gene name	Protein name	Fold change	*p*-value
AKAP12	A-kinase anchor protein 12	1.66	***
CD82	CD82 antigen	1.61	***
GSK3B	Glycogen synthase kinase-3 beta	1.52	*
EPHB2	Ephrin type-B receptor 2	1.5	***
DNMT1	DNA (cytosine-5)-methyltransferase 1	1.48	***
PDCD4	Programmed cell death protein 4	1.42	***
PEBP1	Phosphatidylethanolamine-binding protein 1; hippocampal cholinergic neurostimulating peptide	1.35	***
CDKN1A	Cyclin-dependent kinase inhibitor 1	1.31	**
SFN	14-3-3 protein sigma	1.24	**
STAT3	Signal transducer and activator of transcription 3	1.23	**
GSTP1	Glutathione S-transferase P	1.22	***
LIMD1	LIM domain-containing protein 1	1.2	**

*$P < 0.05$; **$P < 0.01$, and ***$P < 0.001$.

TABLE 2 | Cancer stem cell-related and downregulated proteins by NEDD4 knockdown in MDA-MB-231.

Gene name	Protein name	Fold change	*p*-value
SQSTM1	Sequestosome-1	0.73	***
SUZ12	Polycomb protein SUZ12	0.76	***
UBE2C	Ubiquitin-conjugating enzyme E2 C	0.77	*
EZH2	Histone-lysine N-methyltransferase EZH2	0.81	***
CTGF	Connective tissue growth factor	0.82	*
KIAA1524	Protein CIP2A	0.83	***
YBX1	Nuclease-sensitive element-binding protein 1	0.83	**

*$P < 0.05$; **$P < 0.01$, and ***$P < 0.001$.

RESULTS

The Relationship Between the Expression of NEDD4 and the Survival Rate of Breast Cancer Patients

We assessed the effects of NEDD4 level on the survival rate of breast cancer patients. For that, we used the Kaplan–Meier (KM) plotter, which is a web-based tool that shows the effect of 54,000 genes on the survival rate of cancer patients[2]. The KM plotter revealed that the survival rate of patients with a high expression of NEDD4 was shorter than that of patients with a low expression of NEDD4 in HER2-positive ($HER2^+$) and TNBC, which have been known as relatively aggressive breast cancers (**Figure 1A**). In addition, the statistically significant positive relationship appeared distinctively in the advanced stage of breast cancer as opposed to the early stage (**Figure 1B**). A western blot analysis showed that the expression level of NEDD4 was relatively higher in MDA-MB-231 (TNBC) and MDA-MB-436 (TNBC) than in MCF7 (luminal type) and SK-BR-3 (HER2 + type). As a result, the higher the expression of NEDD4, the more aggressive and the lower the survival rate of the breast cancer.

NEDD4 Knockdown Inhibited the Proliferation and the Migration of MDA-MB-231

For further experiments, the MDA-MB-231 cell line was selected due to its high expression of NEDD4 (**Figure 1C**) and its TNBC properties, including being positive for BCSC markers (30). Proliferation and migration were confirmed in siCONT- and siNEDD4-treated MDA-MB-231. The proliferation was significantly decreased in siNEDD4-treated MDA-MB-231 (**Figure 2A**). The migration was also confirmed to be decreased in knockdown cells by both wound healing assay (**Figure 2B**) and transwell system (**Figure 2C**). The invasiveness was also decreased in siNEDD4-treated MDA-MB-231 than in the control (**Figure 2D**). To make the results more convincing, proliferation, migration, and invasion assays were also conducted in siCONT- and siNEDD4-treated MDA-MB-436, one of the TNBC cell lines. The proliferation was confirmed to be decreased by NEDD4 knockdown by cell counting (**Supplementary Figure 1A**). The migration and the invasion capacity were also decreased in siNEDD4-treated MDA-MB-436 (**Supplementary Figures 1B,C**). Taken together, MDA-MB-231 cell viability, in the form of proliferation and migration, is affected by NEDD4.

Identification and Functional Characterization of Differentially Expressed Proteins in NEDD4-Knocked-Down MDA-MB-231 by LC–MS/MS

To identify the altered proteomes by NEDD4 deletion in MDA-MB-231, LC–MS/MS analysis was conducted. Proteins isolated from siCONT- and siNEDD4-treated MDA-MB-231 were digested using the FASP method and labeled by TMT six-plex reagent. Following this, the combined samples were fractionated to three by styrene divinylbenzene–reversed-phase sulfonate and analyzed by LC–MS/MS (**Figure 3A**). The western blot analysis using anti-NEDD4 antibody showed that the expression of NEDD4 was reduced by more than 90% by transfection of siNEDD4 (**Figure 3B**). The volcano plot showed that 333 proteins were increased and 252 proteins were decreased in siNEDD4-treated MDA-MB-231 as opposed to those in siCONT-treated cells (**Figure 3C** and **Supplementary Tables 2, 3**). Accordingly, the reduced expression of NEDD4 by siNEDD4 treatment was also confirmed by LC–MS/MS.

To further categorize and visualize the function and the pathway enrichment of differentially expressed proteins, gene set enrichment analysis was performed. A previous report determined that poorly differentiated aggressive human breast cancers showed embryonic stem cell-like signatures, including 13 gene sets (31). Interestingly, compared to the control, the NEDD4-knockdowned MDA-MB-231 cells had a lower expression of genes enriched by "MYC_TARGETS_V2 (p-value < 0.05, FDR q-value = 0.23)" and "MYC_TARGETS_V1

[2]https://kmplot.com/analysis/index.php?p=service&cancer=breast

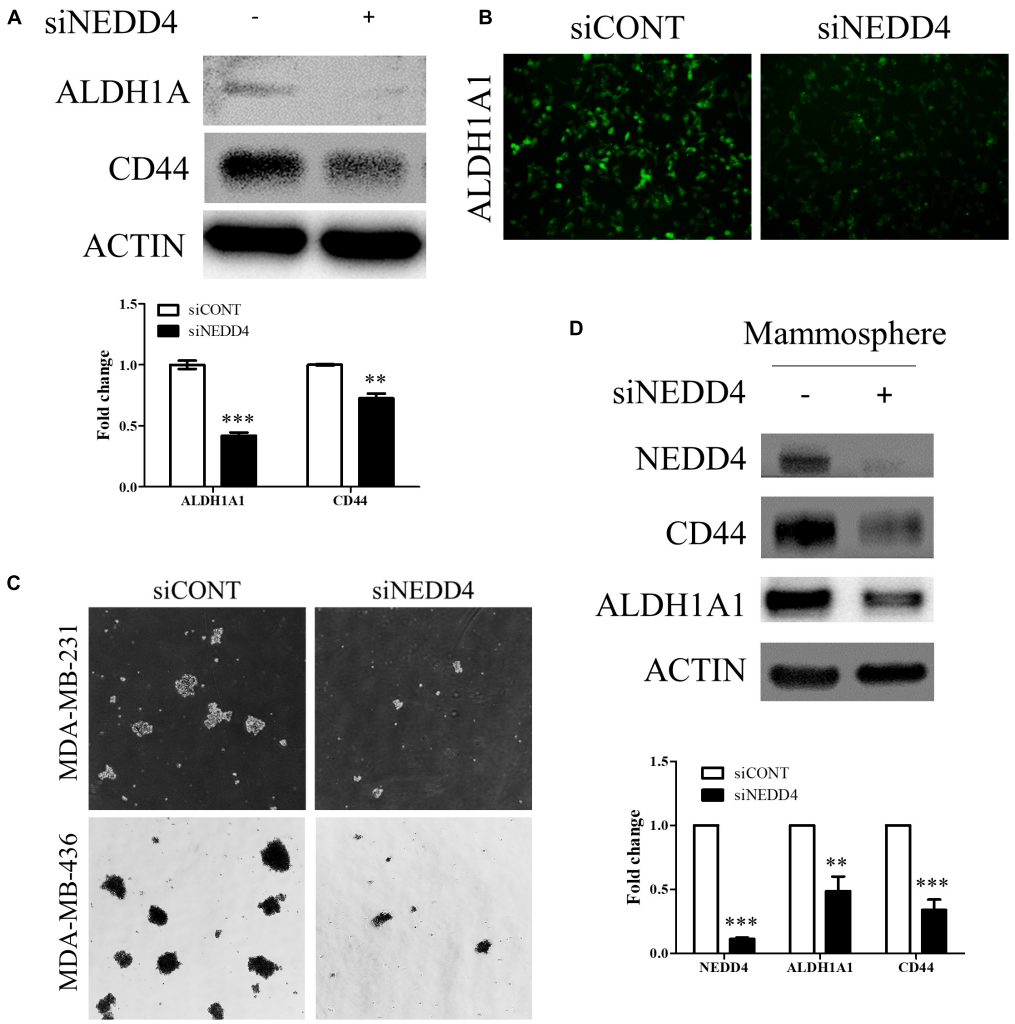


FIGURE 5 | The expression and the activity of breast cancer stem cell (BCSC) markers were decreased in NEDD4-deleted MDA-MB-231. **(A)** The expression of BCSC markers including ALDH1A1 and CD44 was confirmed by western blot in siCONT- and siNEDD4-treated MDA-MB-231. β-actin was used as a loading control. **(B)** The activity of ALDH1A1 was confirmed by the ALDEFLUOR kit system in siCONT/NEDD4-treated MDA-MB-231 cells. The representative images are shown. The average value is shown in the graph (n = 3). **(C)** siCONT/NEDD4-treated MDA-MB-231 cells were induced to mammosphere formation for 2 weeks. The representative images are shown (n = 3). **(D)** The lysates harvested from the mammosphere formation were subjected to western blot using the indicated antibodies. β-actin was used as a loading control. The representative images are shown. The average value is shown in the graph (n = 3) (**P < 0.01 and ***P < 0.001).

(p-value < 0.05, FDR q-value = 0.28)," which are known as embryonic stem cell-like gene signatures (**Figure 3D**).

Next, to distinguish the CSC-related proteins among those differentially expressed by NEDD4 knockdown, up/downregulated proteins were matched with CSC-related proteins excerpted from CSCdb (32). A number of studies have shown that NEDD4 is an oncoprotein that catalyzes the ubiquitination and the degradation of target proteins that are commonly known as tumor suppressors, such as PTEN and LATS (16, 18). We thus hypothesized that upregulated proteins by NEDD4 knockdown include tumor suppressors and potential candidates for ubiquitination targets. In addition, downregulated proteins are regarded as oncoproteins. Therefore, known tumor suppressors (TSGene) (33), and oncogenes (ONGene) (34) were selected from the 333 upregulated and 252 downregulated proteins, respectively. Finally, 12 upregulated proteins, such as AKAP12, CD82, GSK3B, EPHB2, DNMT1, PDCD4, PEBP1, CDKN1A, SFN, STAT3, GSTP1, and LIMD1, were matched with both CSC-related and tumor suppressors (**Figure 4A** and **Table 1**), and seven downregulated proteins, such as SQSTM1, SUZ12, UBE2C, EZH2, CTGF, KIAA1524, and YBX1, were matched with both CSC-related and oncoproteins (**Figure 4B** and **Table 2**). The mRNA expression of some up/downregulated proteins was confirmed by qPCR. In the condition that generally about 30–40% mRNA expression correlates with protein expression, our qPCR data showing that the mRNA expression of CD82 increased (**Figure 4C**) and SUZ12, UBE2C, CTGF, and KIAA1524 decreased (**Figure 4D**) in siNEDD4-treated MAD-MB-231 cells were validated to proteomic data.

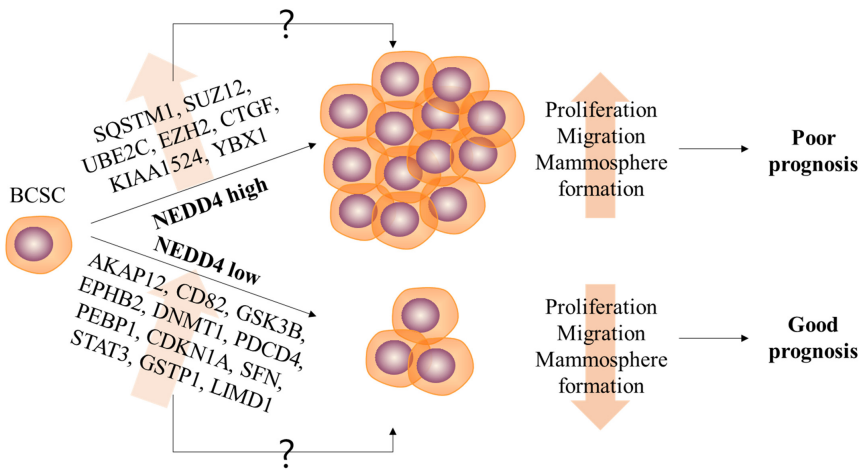


FIGURE 6 | The role of NEDD4 in the regulation of breast cancer stem cell (BCSC) properties in MDA-MB-231. The schematic model shows that NEDD4 is required for proliferation, migration, and stemness maintenance in BCSCs. The breast cancer patients with a high expression of NEDD4 in BCSCs may show a shorter survival period and a poor prognosis due to the increased proliferation, migration, and stemness of BCSCs, while the patients with a low level of NEDD4 in BCSCs may live longer and have a better prognosis. Although detailed researches were needed, our proteomics analysis revealed that the differentially expressed proteins by NEDD4 knockdown were related to BCSCs.

The Properties of BCSC Were Decreased by NEDD4 Knockdown

We also investigated whether the expression and the activity of BCSC markers are altered by NEDD4 knockdown. The expression of ALDH1A1 and CD44 was reduced in siNEDD4-treated MDA-MB-231 cells (**Figure 5A**). The activity of ALDH1A1 was also confirmed with the ALDEFLUOR assay. NEDD4-knocked-down MDA-MB-231 cells resulted in a much weaker green intensity, which indicates low ALDH1A1 activity (**Figure 5B**). These observations led us to test whether NEDD4 knockdown suppresses mammosphere formation, which is also a stem cell-like phenotype. To this end, siCONT- and NEDD4-treated MDA-MB-231 were cultured in mammosphere media for 2 weeks. Expectedly, the knockdown of NEDD4 inhibited big and round mammosphere formation (**Figure 5C**). Also, the expression of BCSC markers including ALDH1A1 and CD44 was decreased by NEDD4 knockdown (**Figure 5D**). Taken together, NEDD4 is required for the maintenance of breast CSC characteristics.

DISCUSSION

The expression and the molecular function of NEDD4 have frequently been studied in various human cancers. However, only a few studies have reported on the role of NEDD4 in CSCs. Targeting CSCs in cancer treatment is very important because of their tumor-initiating and metastatic capacity and since non-CSCs enduring chemotherapy treatment can acquire CSC features leading to poor prognosis (7). In this study, we report the reduced properties of CSCs and CSC-related DEPs by NEDD4 knockdown. The proliferation and the migration of MDA-MB-231 are decreased, and the expression of BCSC markers and the ability to form mammosphere are also significantly reduced by NEDD4 knockdown (**Figure 6**).

Numerous studies have suggested a lot of downstream substrates, and upstream regulators of NEDD4 are known to be involved in carcinogenesis (35, 36). However, only a few studies have investigated the role of NEDD4 in breast cancer (BC) and BCSC. One such study showed that the expression of NEDD4 was elevated, and it promoted the growth of BC (37), yet the BC-type dependency has not been considered in studying the function of NEDD4. In this study, we realized that breast cancer patients with highly expressing NEDD4 show significantly reduced survival rates, specifically in poorly differentiated breast cancer types which are ER-negative breast cancer types, including $HER2^+$ and TNBC (**Figure 1A**). Although the expression and the role of ER in BCSCs remain under debate, several studies have shown that the expression level of ER was very low or none in $CD44^+/CD24^-/ALDH^+$ CSCs (38, 39). Moreover, $ALDH^-$ BCSCs were associated with ER-negativity, HER2-positivity, high histological grade, and poor prognosis (40). Furthermore, breast cancers lacking the expression of ER showed embryonic stem cell-like gene set enrichment (31).

In this study, we focused on the role of NEDD4 in BCSCs, which may account for a shorter survival rate in the patients. Several studies show that the number of BCSCs was increased by stimulation with estradiol in spite of ER-negativity in BCSCs. In ER-negative BCSCs, G-protein coupled receptor (GPCR) 30 could mediate estrogen action. The GPCR-mediated Hippo-TAZ pathway was also required to activate the self-renewal capacity of non-CSCs (41, 42). The increased TAZ activation was detected in poorly differentiated breast cancers, and it also enhanced the metastatic features. Other reports indicate that breast cancers highly expressing type 1 tyrosine kinase-like orphan receptor also have stem cell-like features through activation

of Rho-GTPase and the Hippo-YAP/TAZ signaling pathway (43). NEDD4 has been known as a regulator of the Hippo signaling pathway through ubiquitination and degradation of LATS kinase (18, 19, 44). Although more detailed experiments and analysis are needed, NEDD4 might regulate the stem cell-like characteristics in BCSCs by activating the GPCR-mediated Hippo-YAP/TAZ pathway.

A previous report exhibited that ES gene set signatures including ES exp1, ES exp2, Nanog targets, Oct4 targets, Sox2 targets, NOS targets, NOS TFs, Myc target1, and Myc target2 were overexpressed and gene sets including Suz12 targets, Eed targets, H3K27 bound, and PRC3 targets were underexpressed in ER-negative and high grade 3 breast cancer patients (31). Our TMT-based high-throughput proteomics analysis revealed 585 differentially expressed proteins (DEPs) in siNEDD4-treated MDA-MB-231, and these DEPs are functionally associated with the maintenance of BCSC properties. The two gene sets including Myc_target1 and Myc_target2 are especially enriched in downregulated proteins following NEDD4 knockdown (**Figure 3D**), indicating that the ESC-like features in ER-negative and highly aggressive breast cancers are lost with NEDD4 knockdown.

Our LC–MS/MS analysis also revealed that the expression of both EZH2 and SUZ12 is significantly decreased (**Table 2**), and the mRNA level of SUZ12 is also reduced in siNEDD4-treated MDA-MB-231 (**Figure 4D**). EZH2 and SUZ12 both belong to polycomb-repressive complex 2 (PRC2), which is necessary for the maintenance of stem cell pluripotency and the inhibition of differentiation by activation of histone H3 lysine 27 trimethylation (H3K27me3) that represses the transcription of the differentiation genes (45). It also interacted with BRCA1, a tumor suppressor, in both ES cells and breast cancers and acted as a negative modulator of PRC2. The loss of BRCA1 led to the activation of PRC2, inhibition of cell differentiation, and induction of more aggressive breast cancers (46). PRC2 is composed of four core subunits: EZH2, SUZ12, EED, and RbAp46/48. The expression of EZH2 and SUZ12 was increased in several types of cancers such as prostate, ovarian, and breast cancers, and it initiated tumorigenesis and activated anti-drug responses through preventing the expression of differentiation-related genes and promoting stem cell-like phenotypes (47–50). Thus, our results suggest that NEDD4 is involved in the PRC2 complex which inhibits the transcription of differentiation genes.

In conclusion, our results indicate that the systems which regulate the maintenance of stemness are negatively regulated by NEDD4 knockdown, which affects the formation of the mammosphere and the expression of BCSC markers such as ALDH1A1 and CD44. We believe that such discoveries will help to provide much insight into the development of novel therapeutics and drugs targeting NEDD4 in a variety of CSCs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: http://proteomecentral.proteomexchange.org/cgi/GetDataset, Accession: PXD020585.

AUTHOR CONTRIBUTIONS

S-AJ, DK, and J-YC designed the research, analyzed the data, developed the methodologies, and reviewed and revised the manuscript. S-AJ performed the experiments. S-AJ, DK, and D-BL conducted the bioinformatics analysis. D-BL operated the LC-MS/MS analysis. S-AJ wrote the manuscript draft. J-YC supervised this study. All the authors read and approved the final manuscript.

FUNDING

This research was supported by the Bio and Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (#2016M3A9B6026771).

ACKNOWLEDGMENTS

We would like to thank Johannes Josephus Schabort (Department of Veterinary Biochemistry, Seoul National University, Seoul, South Korea) for English proofreading of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01680/full#supplementary-material

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer.* (2010) 127:2893–917. doi: 10.1002/ijc.25516
2. Hart CD, Migliaccio I, Malorni L, Guarducci C, Biganzoli L, Di Leo A. Challenges in the management of advanced, ER-positive, HER2-negative breast cancer. *Nat Rev Clin Oncol.* (2015) 12:541–52. doi: 10.1038/nrclinonc.2015.99
3. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol.* (1999) 17:2639–48. doi: 10.1200/JCO.1999.17.9.2639
4. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol.* (2008) 26:1275–81. doi: 10.1200/JCO.2007.14.4147
5. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* (2011) 121:2750–67. doi: 10.1172/JCI45014

6. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA*. (2003) 100:3983–8. doi: 10.1073/pnas.0530291100
7. Giovannelli P, Di Donato M, Galasso G, Di Zazzo E, Medici N, Bilancio A, et al. Breast cancer stem cells: the role of sex steroid receptors. *World J Stem Cells.* (2019) 11:594–603. doi: 10.4252/wjsc.v11.i9.594
8. Sin WC, Lim CL. Breast cancer stem cells-from origins to targeted therapy. *Stem Cell Investig.* (2017) 4:96. doi: 10.21037/sci.2017.11.03
9. Idowu MO, Kmieciak M, Dumur C, Burton RS, Grimes MM, Powers CN, et al. CD44(+)/CD24(-/low) cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. *Hum Pathol.* (2012) 43:364–73. doi: 10.1016/j.humpath.2011.05.005
10. O'Conor CJ, Chen T, Gonzalez I, Cao D, Peng Y. Cancer stem cells in triple-negative breast cancer: a potential target and prognostic marker. *Biomark Med.* (2018) 12:813–20. doi: 10.2217/bmm-2017-0398
11. Ingham RJ, Gish G, Pawson T. The Nedd4 family of E3 ubiquitin ligases: functional diversity within a common modular architecture. *Oncogene.* (2004) 23:1972–84. doi: 10.1038/sj.onc.1207436
12. Ye X, Wang L, Shang B, Wang Z, Wei W. NEDD4: a promising target for cancer therapy. *Curr Cancer Drug Targets.* (2014) 14:549–56. doi: 10.2174/1568009614666140725092430
13. Zou X, Levy-Cohen G, Blank M. Molecular functions of NEDD4 E3 ubiquitin ligases in cancer. *Biochim Biophys Acta.* (2015) 1856:91–106. doi: 10.1016/j.bbcan.2015.06.005
14. Dai B, Pieper RO, Li D, Wei P, Liu M, Woo SY, et al. FoxM1B regulates NEDD4-1 expression, leading to cellular transformation and full malignant phenotype in immortalized human astrocytes. *Cancer Res.* (2010) 70:2951–61. doi: 10.1158/0008-5472.CAN-09-3909
15. Weng M, Luo ZL, Wu XL, Zeng WZ. The E3 ubiquitin ligase NEDD4 is translationally upregulated and facilitates pancreatic cancer. *Oncotarget.* (2017) 8:20288–96. doi: 10.18632/oncotarget.15446
16. Wang X, Trotman LC, Koppie T, Alimonti A, Chen Z, Gao Z, et al. NEDD4-1 is a proto-oncogenic ubiquitin ligase for PTEN. *Cell.* (2007) 128:129–39. doi: 10.1016/j.cell.2006.11.039
17. Hong SW, Moon JH, Kim JS, Shin JS, Jung KA, Lee WK, et al. p34 is a novel regulator of the oncogenic behavior of NEDD4-1 and PTEN. *Cell Death Differ.* (2014) 21:146–60. doi: 10.1038/cdd.2013.141
18. Salah Z, Cohen S, Itzhaki E, Aqeilan RI. NEDD4 E3 ligase inhibits the activity of the Hippo pathway by targeting LATS1 for degradation. *Cell Cycle.* (2013) 12:3817–23. doi: 10.4161/cc.26672
19. Jeon SA, Kim DW, Cho JY. Neural precursor cell-expressed, developmentally down-regulated 4 (NEDD4) regulates hydrogen peroxide-induced cell proliferation and death through inhibition of Hippo signaling. *FASEB J.* (2019) 33:14772–83. doi: 10.1096/fj.201901404R
20. Xu C, Fan CD, Wang X. Regulation of Mdm2 protein stability and the p53 response by NEDD4-1 E3 ligase. *Oncogene.* (2015) 34:281–9. doi: 10.1038/onc.2013.557
21. Wilkin MB, Carbery AM, Fostier M, Aslam H, Mazaleyrat SL, Higgs J, et al. Regulation of notch endosomal sorting and signaling by Drosophila Nedd4 family proteins. *Curr Biol.* (2004) 14:2237–44. doi: 10.1016/j.cub.2004.11.030
22. Cao XR, Lill NL, Boase N, Shi PP, Croucher DR, Shan H, et al. Nedd4 controls animal growth by regulating IGF-1 signaling. *Sci Signal.* (2008) 1:ra5. doi: 10.1126/scisignal.1160940
23. Fan CD, Lum MA, Xu C, Black JD, Wang X. Ubiquitin-dependent regulation of phospho-AKT dynamics by the ubiquitin E3 ligase, NEDD4-1, in the insulin-like growth factor-1 response. *J Biol Chem.* (2013) 288:1674–84. doi: 10.1074/jbc.M112.416339
24. Fang DF, He K, Wang N, Sang ZH, Qiu X, Xu G, et al. NEDD4 ubiquitinates TRAF3 to promote CD40-mediated AKT activation. *Nat Commun.* (2014) 5:4513. doi: 10.1038/ncomms5513
25. Chen C, Matesic LE. The Nedd4-like family of E3 ubiquitin ligases and cancer. *Cancer Metastasis Rev.* (2007) 26:587–604. doi: 10.1007/s10555-007-9091-x
26. Jeon SA, Lee JH, Kim DW, Cho JY. E3-ubiquitin ligase NEDD4 enhances bone formation by removing TGFbeta1-induced pSMAD1 in immature osteoblast. *Bone.* (2018) 116:248–58. doi: 10.1016/j.bone.2018.08.012
27. Cox J, Matic I, Hilger M, Nagaraj N, Selbach M, Olsen JV, et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc.* (2009) 4:698–705. doi: 10.1038/nprot.2009.36
28. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem.* (2003) 75:4646–58. doi: 10.1021/ac0341261
29. Oberg AL, Mahoney DW, Eckel-Passow JE, Malone CJ, Wolfinger RD, Hill EG, et al. Statistical analysis of relative labeled mass spectrometry data from complex samples using ANOVA. *J Proteome Res.* (2008) 7:225–33. doi: 10.1021/pr700734f
30. Li W, Ma H, Zhang J, Zhu L, Wang C, Yang Y. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. *Sci Rep.* (2017) 7:13856. doi: 10.1038/s41598-017-14364-2
31. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet.* (2008) 40:499–507. doi: 10.1038/ng.127
32. Shen Y, Yao H, Li A, Wang M. CSCdb: a cancer stem cells portal for markers, related genes and functional information. *Database (Oxford).* (2016) 2016:baw023. doi: 10.1093/database/baw023
33. Zhao M, Kim P, Mitra R, Zhao J, Zhao Z. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res.* (2016) 44:D1023–31. doi: 10.1093/nar/gkv1268
34. Liu Y, Sun J, Zhao M. ONGene: a literature-based database for human oncogenes. *J Genet Genomics.* (2017) 44:119–21. doi: 10.1016/j.jgg.2016.12.004
35. Huang X, Chen J, Cao W, Yang L, Chen Q, He J, et al. The many substrates and functions of NEDD4-1. *Cell Death Dis.* (2019) 10:904. doi: 10.1038/s41419-019-2142-8
36. Wang ZW, Hu X, Ye M, Lin M, Chu M, Shen X. NEDD4 E3 ligase: Functions and mechanism in human cancer. *Semin Cancer Biol.* (2020). doi: 10.1016/j.semcancer.2020.03.006
37. Wan L, Liu T, Hong Z, Pan Y, Sizemore ST, Zhang J, et al. NEDD4 expression is associated with breast cancer progression and is predictive of a poor prognosis. *Breast Cancer Res.* (2019) 21:148. doi: 10.1186/s13058-019-1236-7
38. Morimoto K, Kim SJ, Tanei T, Shimazu K, Tanji Y, Taguchi T, et al. Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression. *Cancer Sci.* (2009) 100:1062–8. doi: 10.1111/j.1349-7006.2009.01151.x
39. Asselin-Labat ML, Vaillant F, Sheridan JM, Pal B, Wu D, Simpson ER, et al. Control of mammary stem cell function by steroid hormone signalling. *Nature.* (2010) 465:798–802. doi: 10.1038/nature09027
40. Ginestier C, Korkaya H, Dontu G, Birnbaum D, Wicha MS, Charafe-Jauffret E. [The cancer stem cell: the breast cancer driver]. *Med Sci (Paris).* (2007) 23:1133–9. doi: 10.1051/medsci/200723121133
41. Cordenonsi M, Zanconato F, Azzolin L, Forcato M, Rosato A, Frasson C, et al. The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell.* (2011) 147:759–72. doi: 10.1016/j.cell.2011.09.048
42. Bartucci M, Dattilo R, Moriconi C, Pagliuca A, Mottolese M, Federici G, et al. TAZ is required for metastatic activity and chemoresistance of breast cancer stem cells. *Oncogene.* (2015) 34:681–90. doi: 10.1038/onc.2014.5
43. Zhang S, Zhang H, Ghia EM, Huang J, Wu L, Zhang J, et al. Inhibition of chemotherapy resistant breast cancer stem cells by a ROR1 specific antibody. *Proc Natl Acad Sci USA.* (2019) 116:1370–77. doi: 10.1073/pnas.1816262116
44. Bae SJ, Kim M, Kim SH, Kwon YE, Lee JH, Kim J, et al. NEDD4 controls intestinal stem cell homeostasis by regulating the Hippo signalling pathway. *Nat Commun.* (2015) 6:6314. doi: 10.1038/ncomms7314
45. Holm K, Grabau D, Lovgren K, Aradottir S, Gruvberger-Saal S, Howlin J, et al. Global H3K27 trimethylation and EZH2 abundance in breast tumor subtypes. *Mol Oncol.* (2012) 6:494–506. doi: 10.1016/j.molonc.2012.06.002
46. Wang L, Zeng X, Chen S, Ding L, Zhong J, Zhao JC, et al. BRCA1 is a negative modulator of the PRC2 complex. *EMBO J.* (2013) 32:1584–97. doi: 10.1038/emboj.2013.95
47. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature.* (2002) 419:624–9. doi: 10.1038/nature01075

48. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc Natl Acad Sci USA*. (2003) 100:11606–11. doi: 10.1073/pnas.1933744100
49. Li H, Cai Q, Godwin AK, Zhang R. Enhancer of zeste homolog 2 promotes the proliferation and invasion of epithelial ovarian cancer cells. *Mol Cancer Res.* (2010) 8:1610–8. doi: 10.1158/1541-7786.MCR-10-0398
50. Li H, Cai Q, Wu H, Vathipadiekal V, Dobbin ZC, Li T, et al. SUZ12 promotes human epithelial ovarian cancer by suppressing apoptosis via silencing HRK. *Mol Cancer Res.* (2012) 10:1462–72. doi: 10.1158/1541-7786.MCR-12-0335

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 © 2020 Jeon, Kim, Lee and Cho. 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 Oncology

METHODS
published: 23 September 2020
doi: 10.3389/fonc.2020.543997

An Integrated Quantitative Proteomics Workflow for Cancer Biomarker Discovery and Validation in Plasma

*Vipin Kumar, Sandipan Ray†, Saicharan Ghantasala and Sanjeeva Srivastava**

Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India

OPEN ACCESS

Edited by:
Suman S. Thakur, Center for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Jean-Charles Sanchez, Université de Genève, Switzerland
Paul Townsend, The University of Manchester, United Kingdom
Edouard Nice, Monash University, Australia

***Correspondence:**
Sanjeeva Srivastava
sanjeeva@iitb.ac.in

†Present address:
Sandipan Ray, Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 19 March 2020
Accepted: 17 August 2020
Published: 23 September 2020

Citation:
Kumar V, Ray S, Ghantasala S and Srivastava S (2020) An Integrated Quantitative Proteomics Workflow for Cancer Biomarker Discovery and Validation in Plasma. Front. Oncol. 10:543997. doi: 10.3389/fonc.2020.543997

Blood plasma is one of the most widely used samples for cancer biomarker discovery research as well as clinical investigations for diagnostic and therapeutic purposes. However, the plasma proteome is extremely complex due to its wide dynamic range of protein concentrations and the presence of high-abundance proteins. Here we have described an optimized, integrated quantitative proteomics pipeline combining the label-free and multiplexed-labeling-based (iTRAQ and TMT) plasma proteome profiling methods for biomarker discovery, followed by the targeted approaches for validation of the identified potential marker proteins. In this workflow, the targeted quantitation of proteins is carried out by multiple-reaction monitoring (MRM) and parallel-reaction monitoring (PRM) mass spectrometry. Thus, our approach enables both unbiased screenings of biomarkers and their subsequent selective validation in human plasma. The overall procedure takes only ∼2 days to complete, including the time for data acquisition (excluding database searching). This protocol is quick, flexible, and eliminates the need for a separate immunoassay-based validation workflow in blood cancer biomarker investigations. We anticipate that this plasma proteomics workflow will help to accelerate the cancer biomarker discovery program and provide a valuable resource to the cancer research community.

Keywords: cancer biomarker, multiplexed quantitative proteomics, targeted proteomics, label-free quantitation, multiple reaction monitoring, parallel reaction monitoring

INTRODUCTION

Plasma is an attractive and reliable sample for cancer research due to its easy accessibility, and plasma proteome can provide a plethora of important information regarding the normal physiological states as well as the cancer-induced alterations in our body (1, 2). Importantly, recent studies showed whole blood as a specimen for liquid biopsy for personalized medicine applications and monitoring the therapeutic responses to the treatment of cancers (3, 4). Mass spectrometry (MS)—based label-free and multiplexed label-based proteomics profiling of the plasma or serum proteome is widely used for unbiased discovery of potential biomarkers for diverse types of human diseases including cancers, infectious diseases, cardiovascular and metabolic disorders (2, 5–8).

In recent years, multiple-reaction monitoring (MRM) and parallel-reaction monitoring (PRM) mass spectrometry approaches have emerged as attractive alternatives for protein immunoassays (9). These targeted proteomics approaches can accurately measure concentrations of multiple

proteins in complex biological samples, such as plasma (10–13). Importantly, results obtained in multiplexed MRM/PRM-MS assays correlate well with immunoassay-based measurements (10, 14). One key advantage of these targeted MS-assays is that these allow quantification of variants and modified forms of the proteins by targeting their specific peptide sequences (15, 16). Quantification by traditional immunoassay-based techniques such as Western blotting is based on a single antibody that is often inadequately characterized and protein quantification solely depends on a single signal (17). On the contrary, the quality of the isotopically labeled reference peptides used in MRM or PRM-based methods could be easily evaluated by a fragment ion spectrum and these approaches use multiple signals for obtaining more reliable and robust quantification (17). Moreover, immunoassay-based techniques are often difficult to perform for multiple targets due to the low-throughput of these approaches and the unavailability of suitable antibodies for many proteins. To this end, MRM and PRM-based approaches allow the accurate quantification of hundreds of peptides in a single injection/run of mass spectrometry and therefore are more high-throughput compared to the conventional immunoassay-based measurements. Consequently, a combined workflow involving both discovery and validation phase quantitative proteomics techniques would be extremely beneficial for cancer biomarker research.

There are several methods describing sample processing for quantitative proteomics analysis of plasma samples in various cancers, while we have demonstrated here a combined method for both discovery and validation of protein markers in plasma samples. In this respect, we have extensive experience of applying label-based multiplexed quantitative proteomics for the discovery of biomarker panels in cancer and other diseases (18–24). Such multiplexing using stable isotope labeling results in increased throughput, higher precision, better reproducibility, reduced technical variations and fewer missing values (8, 20, 25–30). Further, proteomic profiling using label-free quantitation (LFQ) is another attractive method for cancer biomarker quantification (23, 31). In recent years, we have reported targeted quantitation of proteins by Multiple Reaction Monitoring (MRM) mass spectrometry (18, 32). Here, we have described an amalgamated analysis pipeline for plasma biomarker analysis by integrating the know-how of different quantitative & targeted proteomics methods (**Figures 1A–D**).

EXPERIMENTAL DESIGN

In this integrated quantitative proteomics pipeline, three biological pool of plasma samples were analyzed for obtaining a comprehensive proteome profile, and subsequent validation of a few selective peptides. Each of the three plasma pools (named as samples A, B, and C) was a uniform mixture of ten different individual plasma samples. In order to perform targeted proteomics analyses, a pool of 21 heavy labeled synthetic peptides were spiked into the plasma samples at a different ratio. Global quantitative proteomics was performed using both label-free and label-based such as Isobaric tags for relative and absolute quantitation (iTRAQ 4-plex) and Tandem Mass Tag™ (TMT 6-plex) quantitation approaches (**Figures 1B,C**), while the targeted proteomics was carried out using MRM and PRM-based MS assays (**Figure 1D**). In iTRAQ experiment, we have used different amount of digested peptides per label to determine the minimum amount of peptides to be labeled and the accuracy of the quantitation.

This protocol consists of label-free and label-based (iTRAQ and TMT) proteome profiling methods for cancer biomarker discovery, followed by the targeted approaches (MRM and PRM) for validation of a few potential marker proteins.

STEPWISE PROCEDURE

Plasma Sample Preparation Timing 20 min

1. Collect the blood samples into anticoagulant-treated tubes e.g., EDTA-treated or citrate-treated tubes.
   CRITICAL: Avoid the use of heparin tubes, heparin can often be contaminated with endotoxin, which can stimulate white blood cells to release cytokines.
2. Remove the cells from blood by centrifugation for 15 min at 2,000 ×g, and the resulting supernatant will be plasma.
   CRITICAL STEP: The temperature should be maintained at 2–8°C while handling the samples.
   PAUSE POINT: The samples can be aliquoted in 0.5 ml tubes and stored in −80°C for long-term storage (6–8 months).

Depletion of High Abundant Proteins Timing 1 h

3. Equilibrate the depletion spin column at room temperature (room temperature is 25°C).
4. Remove the column screw cap and add 15 µl of crude plasma sample directly to the resin slurry in the column.
   CRITICAL STEP: Ensure resin slurry is not dried, and the protein concentration of the plasma sample is around 50–60 µg/µl.
5. Cap the column and invert the column several times until the resin is completely suspended in the solution.
6. Incubate the mixture in the column with gentle end-to-end mixing for 60 min at room temperature. Alternatively, vortex every 5 min.
   CRITICAL STEP: Make sure the sample mixes with the resin during the incubation period.
7. Twist off the bottom closure and loosen the cap. Place column into a 2 ml collection tube and centrifuge at 1000 × g for 2 min.
8. Discard the column containing the resin.
9. The filtrate contains depleted plasma (vol. 300 µl approx.) with the top 12 proteins removed.

Protein Quantification and Sample Preparation for Digestion Timing 2 h

10. Reduce the volume of plasma samples up to 75–100 µl using vacuum centrifuge and quantify using Bradford's reagent following the manufacturer's instruction.
11. Check the quality of depleted plasma samples by running on SDS-PAGE and take 50 µg of proteins and dry it completely.
12. Denature the plasma samples by adding 10 µl of 6 M urea.

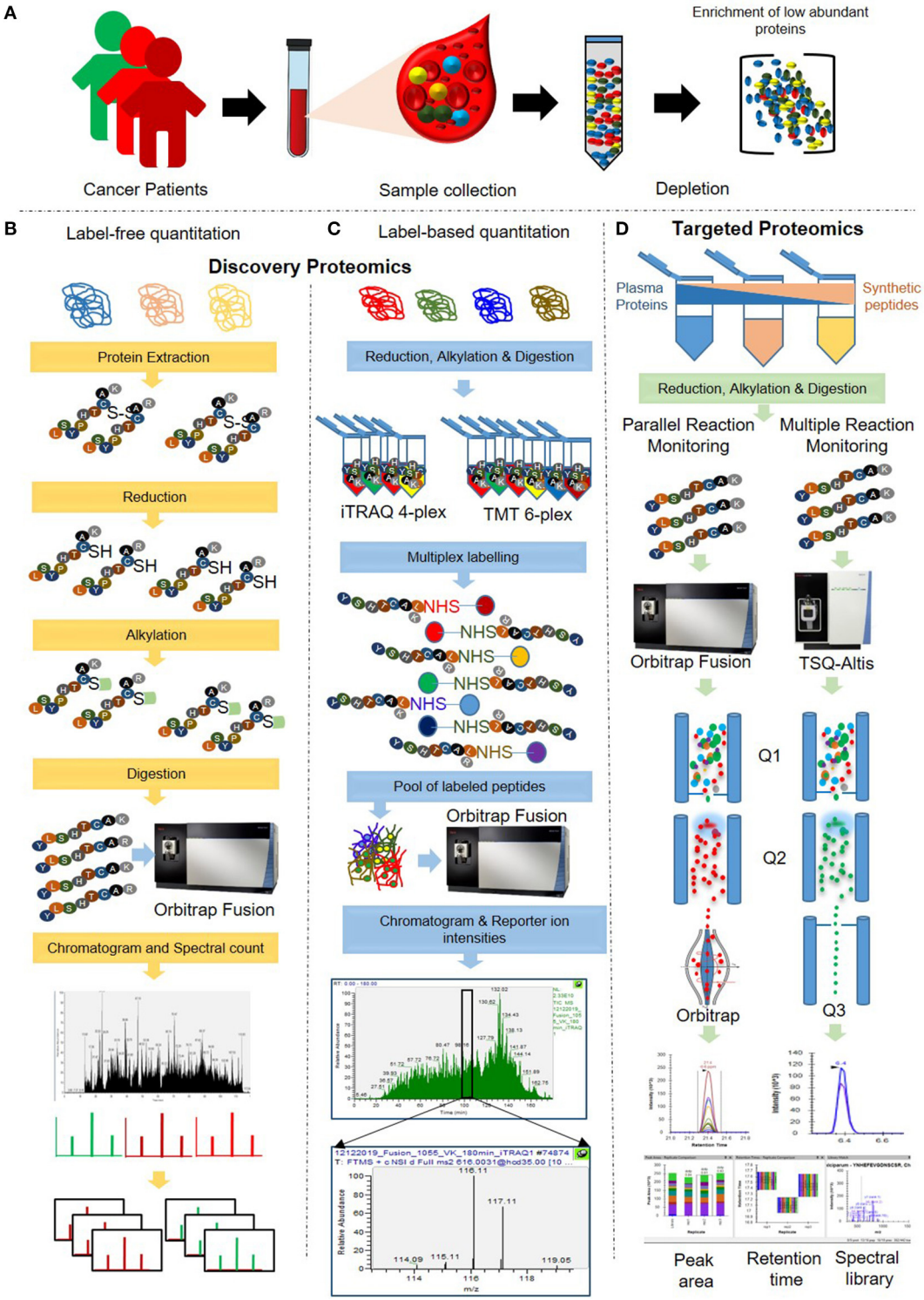


FIGURE 1 | Schematic representation of the integrated workflow for cancer biomarker discovery and validation in plasma. **(A)** Crude plasma samples were depleted using depletion columns for the removal of the top 12 highly abundant proteins. **(B)** Depleted plasma samples were subjected to in-solution digestion, and MS analysis was performed using a label-free quantitation approach. **(C)** The digested peptides were labeled using iTRAQ/TMT reagents and subjected to MS analysis for label-based quantitation. **(D)** Samples were prepared by spiking with the varying amounts of heavy labeled synthetic peptides. Targeted quantification of the spiked synthetic peptides and a few selected potential cancer markers was carried out using Multiple Reaction Monitoring (MRM) and Parallel Reaction Monitoring approach (PRM) approaches.

13. Reduce disulfide bonds by adding tris (2-carboxyethyl) phosphine (TCEP) to a final concentration of 20 mM. Incubate the sample at 37°C for 60 min.
14. Alkylate reduced cysteine residues by adding iodoacetamide (IAA) to a final concentration of 40 mM. Incubate at room temperature (RT) in the dark for 30 min.

Enzymatic Digestion of Plasma Proteins Timing 6–8 h

15. Further, dilute the urea concentration by adding 50 mM ammonium bicarbonate in a 1:6 ratios.
    CRITICAL STEP: The urea concentration should be <1 M. If you are using trypsin as a digestion enzyme.
16. Add trypsin (Pierce) at an enzyme/substrate ratio of 1:50 and incubate at 37°C with shaking on a table-top shaker set at 500 rpm for 6–8 h.
17. Stop the digestion with formic acid (FA) to a final concentration of 1%.

Desalting of the Digested Peptides Timing 1 h per 5–10 Samples

18. Prepare the C18 desalting column by using Empore C18 extraction disks. Pack one plug of C18 material into each stage tip (200 µl pipette tips) for a total binding capacity of ∼25 µg total. Create extraction disks using 200 µl tips, as shown in **Figure S1**.
19. Activate the desalting column with 50 µl of methanol. Centrifuge at 1,000 g for 2 min at RT and discard the liquid from the collection vial. Repeat this step two times.
    CRITICAL: All subsequent centrifugation steps for desalting are for the same duration at the same speed and RT.
20. Wash the desalting column with 50 µl of acetonitrile. Centrifuge at 1,000 g for 2 min at RT and discard the liquid from the collection vial. Repeat this step two times.
21. Equilibrate the desalting column twice with 50 µl of 0.1% (v/v) FA (solvent A). Repeat the centrifugation step. Discard the liquid.
22. Reconstitute the samples in 50 µl of 0.1% (v/v) FA and vortex for 10 min. at slow speed.
23. Pass the reconstituted samples through the desalting column using either syringe or centrifuge at 1,000 g for 2 min. Repeat this step at least 4 times.
    CRITICAL: Ensure that there is no trapping of the air bubble in the desalting column.
24. Store the flow-through at 4°C.
25. Wash the samples twice with 50 µl of solvent A. Repeat the centrifugation step. Discard the liquid.
26. Pass 50 µl of 40% (v/v) acetonitrile in 0.1% (v/v) FA and collect the eluate in new 1.5 ml screw cap. Repeat this step with 50% and 60% (v/v) acetonitrile in 0.1% (v/v) FA and collect the eluate in same vial.
27. Dry 150 µl of desalted sample using a speed vacuum centrifuge.
    PAUSE POINT: Store the desalted peptides at−20°C till further process.

Quantification of the Desalted Peptides Timing 10 min

28. Reconstitute the desalted peptides in 0.1% (v/v) FA.
29. Wipe the µDrop plate with 70% ethanol using lint-free tissue papers.
    CRITICAL: Avoid using normal tissue paper and 70% isopropanol.
30. Blank to be used is 0.1% FA (2 µl).
31. Spot 2 µl of samples onto the µDrop plate along with the blank.
32. Click plate layout → Select µDrop plate from the dropdown menu → Add details about the plate map
33. Click protocol → Absorbance → Multiple wavelengths → Add 205 and 280 nm.
34. Click results → blank subtraction
35. Run plate out → Place the µDrop plate in designated position → Run plate in → Start
36. Calculate E^{205} using the following formula: $E^{205} = 27/1\text{-}3.85{*}A^{280}/A^{205}$.
    CRITICAL: The value of E205 should be 31 ± 3 $mLmg^{-1}cm^{-1}$. If value E205 is not lying in this range, the sample may not be properly digested.
37. Calculate conc. using formula: Absorbance (A^{205}) = E^{205*} conc. * path length (0.05)
38. Conc. (µg/µl) = A^{205}/ E^{205} (calculated from above formula) * path length (0.05)

EXPERIMENT 1: LABEL-FREE QUANTITATION (LFQ) OF PLASMA SAMPLE TIMING 2 h PER SAMPLE

39. The desalted peptides can be run for label-free quantitation using the below-mentioned LC (Section A) and MS (Section B) parameters. We observed good reproducibility between three technical replicates (see anticipated results below).

A. LC Parameters

i. Take 2 µg of digested peptides and make up the volume to 10 µl.
ii. **CRITICAL:** The concentration of desalted peptides will be 200 ng/µl.
iii. Place the vials in the auto-sampler stand of nLC 1200.
iv. Equilibrate the pre-column (Thermo Fisher Scientific, P/N 164564, S/N 10694527) and analytical column (Thermo Fisher Scientific, P/N ES803A, S/N 10918620) five times of column volume with 0.1% (v/v) FA.
v. Load 1 µg of digested peptides onto the column using the nLC 1200 system.
vi. Set up the LC gradient based on sample complexity. We have used 120 min LC gradient for label-free quantitation of the plasma samples. The brief details of LC gradient are mentioned in **Supplementary Method 1**, **2**.

B. MS Parameters

i. Open Thermo Scientific Xcalibur software double click on instrument setup and select the template from peptides-ID with default parameters (**Figure S2**).

ii. Populate the MS parameters from **Figure S3** and save as a new method (**Supplementary Method 3**).
iii. Open Thermo Scientific Xcalibur software, double click on sequence setup and fill sample details such as sample type, sample name, file save location, instrument method file, the volume of injection, and position of the sample.
iv. Select the row and click on the run sequence.

EXPERIMENT 2: LABEL-BASED QUANTIFICATION (ITRAQ 4-PLEX/ TMT 6-PLEX) OF PLASMA SAMPLE TIMING 5 h

40. The digested peptides can be labeled with iTRAQ, TMT reagents, etc. for label-based quantification. We have used iTRAQ 4-plex and TMT 6-plex for the labeling of digested peptides. The procedure for labeling and parameters of LC and MS is mentioned below.

Labeling of Digested Peptides Using iTRAQ Reagents

i. Allow each vial of iTRAQ® reagent required to reach room temperature (∼5 min).
ii. Spin each vial (30 s) to bring the solution to the bottom of the vial.
iii. **CRITICAL:** Please check the vial. There should be a 10–15 μl solution.
iv. Add 70 μl of ethanol to each iTRAQ® Reagent vial.
v. Vortex each vial to mix (30 s), and then spin (10 s).
vi. Transfer the contents of the iTRAQ® reagent vial to each sample tube (114, 115, 116, and 117). In this experiment, 12.5, 25, 50, and 100 μg of digested peptides labeled with the content of each vial's iTRAQ reagents 114, 115, 116, and 117, respectively.
vii. **CRITICAL STEP:** Organic part of the mixture should be >70%; if not, add more ethanol to keep it above 70%.
viii. Vortex each tube to mix (30 s), and then spin (10 s).
ix. **CRITICAL**: Check the pH by placing 1 μl of the solution on pH paper with a pH range of 8.0 to 10.0. If necessary, add up to 10 μl of dissolution buffer—plasma to adjust the pH to >8.
x. Incubate at room temperature for 90 min.
xi. Add 100 μl of Milli-Q water to quench the reaction.
xii. **CRITICAL**: Ensure the aqueous part of the mixture >30%.
xiii. Incubate the tubes at room temperature for 30 min.
xiv. Combine the contents of all iTRAQ Reagent-labeled sample tubes into one tube.
xv. Vortex each tube to mix (30 s), and then spin (10 s).
xvi. Dry the tube containing all the combined iTRAQ labeled peptides.

LC Parameters

i. Follow the steps 39Ai-39Avi from #experiment 1.
ii. Set up the LC gradient based on sample complexity. We have used a 180 min LC gradient for label-based quantitation (iTRAQ) of a plasma sample. The brief details are mentioned in the reagent set up.

Generate an Instrument Method for iTRAQ Technique

i. The MS parameters for label-based quantitation were the same, which were used for label-free quantitation except for collision energy. In the case of label-based quantitation, 35% collision energy was used for MS/MS fragmentation (**Supplementary Method 3**).

Labeling of Digested Peptides Using TMT 6-Plex Reagents and Fractionation Using High-pH Reverse-Phase Technique

i. Allow each vial of TMT 6-plex reagent to reach room temperature (∼5 min).
ii. Spin each vial (2 min) to bring the solution to the bottom of the vial by occasional vortexing.
iii. Add 45 μl of anhydrous acetonitrile to each TMT 6-plex Reagent vial.
iv. Vortex each vial to mix (30 s), and then spin (10 s).
v. Carefully add 40 μl of the TMT label reagent to each 50 μg of digested proteins.
vi. Vortex each tube to mix (30 s), and then spin (10 s).
vii. Incubate the reaction at room temperature for 90 min.
viii. Add 2 μl of 5% hydroxylamine to the sample and incubate for 30 min to quench the reaction.
ix. Combine the contents of all TMT reagent -labeled sample tubes into one tube.
x. Vortex each tube to mix (30 s), and then spin (10 s).
xi. Dry the tube containing all the combined TMT labeled peptides.
xii. Fractionate the labeled peptide samples following manufacturer's instructions [Pierce™ High pH Reversed-Phase Peptide Fractionation Kit (Thermo Scientific™, cat no. 84868)].

LC Parameters

i. Follow the steps 39Ai-39Avi from experiment #1.
ii. Set up the LC gradient based on sample complexity. We have used a 90 min LC gradient for label-based quantitation (TMT 6-plex) of a plasma sample. The brief details of LC gradient are mentioned in **Supplementary Method 1**, **2**.

Generate an Instrument Method for TMT 6-plex Technique

i. The MS parameters for label-based quantitation were the same, which were used for label-free quantitation except for collision energy. In the case of label-based quantitation, 35% collision energy was used for MS/MS fragmentation (**Supplementary Method 3**).

EXPERIMENT 3: MULTIPLE REACTION MONITORING (MRM) ASSAY TIMING 1.5 h

41. The MRM assay was optimized using 21 heavy label synthetic peptides and then endogenous peptides were monitored using the optimized MRM protocol. The steps

for method generation and parameters of LC and MS are mentioned below.

Instrument Method Generation for MRM Using Skyline

i. Load the sequence of synthetic peptides into Skyline and set the parameters for peptide and transition setting as mentioned in **Figures S4**, **S5**.
ii. Export the unscheduled transition list as a single method from Skyline (**Figure S6A**).
iii. Import the unscheduled transition list as an Inclusion list in a MRM acquisition method in Xcalibur.

LC Parameters

i. Follow the steps 39Ai-39Avi from #experiment.
ii. We have used 60 min LC gradient for Multiple Reaction Monitoring (MRM) of a plasma sample. The brief details of LC gradient are mentioned in **Supplementary Methods 1**, **2**.

Set Up Instrument Method for MRM

i. Open Thermo Scientific Xcalibur software double click on instrument setup and select the template from the MRM template with default parameters.
ii. Import the unscheduled transition list as an Inclusion list in a MRM acquisition method in Xcalibur.
iii. Populate the MS parameters from **Figure S7** and save as a new method.
iv. Open Thermo Scientific Xcalibur software double click on sequence setup and fill sample details such as sample type, sample name, file save location, instrument method file, the volume of injection, and position of the sample.
v. Select the row and click on the run sequence.

EXPERIMENT 4: PARALLEL REACTION MONITORING (PRM) ASSAY TIMING 1.5 h

42. The PRM assay was optimized using 21 heavy label synthetic peptides, and then endogenous peptides were monitored using the optimized protocol. The steps for method generation and parameters of LC and MS are mentioned below.

Instrument Method Generation for PRM Using Skyline

i. Load the sequence of synthetic peptides into Skyline and set the parameters as mentioned in **Figures S4**, **S5**.
ii. Export the unscheduled isolation list as a single method from Skyline (**Figure S6B**).

LC Parameters

i. Follow the steps 39Ai-39Avi from experiment #1.
ii. We have used the same LC gradient for PRM, which we have used for MRM.

Set Up Instrument Method for PRM

i. Open Thermo Scientific Xcalibur software double click on instrument setup and select the template from MS^n with default parameters.
ii. Import the unscheduled isolation list as an Inclusion list in a Targeted-MS^2 acquisition method in Xcalibur and populate the MS parameters from **Figure S8** and save as a new method.
iii. Open Thermo Scientific Xcalibur software, double click on sequence setup and fill sample details such as sample type, sample name, file save location, instrument method file, the volume of injection, and position of the sample.
iv. Select the row and click on the run sequence.

DATA ANALYSIS TIMING AROUND 1 D

43. The proteomic data analysis of global and targeted experiments performed using different tools.

Global Proteomics Data Analysis

i. Raw instrument files were processed using Proteome Discoverer (PD) version 2.2 (Thermo Fisher Scientific). MS2 spectra were searched using the Sequest HT and Mascot (v2.6.0) search engine against *Homo sapiens* fasta (74,212 sequence entries, dated: 22/08/2019,) from Uniprot database (Proteome ID: UP000005640, Organism ID: 9606). All searches were configured with dynamic modifications for the iTRAQ reagents (+144.102 Da) on lysine and N-termini, and for TMT reagents (+229.163 Da) on lysine and N-termini of the peptide and oxidation of methionine residues (+15.9949 Da) and static modification as carbamidomethyl (+57.021 Da) on cysteine, monoisotopic masses, and trypsin cleavage (max 2 missed cleavages). The peptide precursor mass tolerance was 10 ppm, and MS/MS tolerance was 0.05 Da. The false discovery rate (FDR) for proteins, peptides, and peptide spectral matches (PSMs) peptides were kept 1%. The quantification values for proteins were exported from proteome discoverer 2.2. The brief parameters were mentioned in **Table 1**. The .raw files from the label-free method were searched against the same database. Most of the proteome discoverer parameters were kept the same as above mentioned for iTRAQ 4-plex method except dynamic modifications for the iTRAQ reagents (+144.102 Da) on lysines and N-termini of a peptide and for TMT reagents (+229.163 Da) on lysine and N-termini of the peptide.
ii. We normalized the data sets using the abundance of total peptide for the identification of differentially expressed proteins. The normalization by total peptide amount is the default option in Proteome Discoverer (v2.2). In this case, it sums the peptide group abundances for each sample and determines the maximum sum for all files, and it calculates the normalization factor using the sum of the sample and the maximum sum in all files.

The users may also use additional data normalization in subsequent steps. There are several normalization approaches, including central tendency, linear regression, locally weighted regression, quantile techniques, and others (33). The normalization methods are evaluated in terms of their ability to reduce variation between technical replicates. Although all these methods can reduce the systematic bias to some extent, each approach has its own advantages and disadvantages (33–35). Therefore, the selection of the normalization

TABLE 1 | The brief details of proteome discoverer parameters.

Description	Parameters	Description	Parameters
Processing Workflow		**Consensus Workflow**	
Spectrum Files RC		Feature mapper	
Precursor Mass Tolerance	20 ppm	Maximum RT shift	5
Fragment mass tolerance	0.05 Da	Mass tolerance	10
Enzyme name	Trypsin	Peptide validator	
Mascot		Validation mode	Only PSM level FDR calculation base
Enzyme name	Trypsin	Target FDR (Strict) for PSMs	0.05
Maximum missed cleavage	2	Target FDR (Relaxed) for PSMs	0.01
Precursor Mass Tolerance	10 ppm	Target FDR (Strict) for Peptide	0.01
Fragment mass tolerance	0.05 Da	Target FDR (Relaxed) for Peptide	0.05
Dynamic modification	**LFQ:** Oxidation (M), Phospho (STY); **iTRAQ:** iTRAQ 4plex, Oxidation (M), Phospho (STY); **TMT:** TMT 6-plex, Oxidation (M), Phospho (STY)	Validation based on	q-value
Static modification	Carboamidomethyl (C)	Use concatenated	TRUE
Sequest		Precursor ions quantifier	
Enzyme name	Trypsin	Normalization mode	Total peptide
Maximum missed cleavage	2	Scaling mode	On all average
Precursor Mass Tolerance	10 ppm	Imputation mode	Low abundance resampling
Fragment mass tolerance	0.05 Da	Protein marker	
Dynamic modification	**LFQ:** Oxidation (M), Acetyl (N-Terminus), Phospho (STY); **iTRAQ:** iTRAQ 4plex (K and N-Terminal), Oxidation (M), Acetyl (N-Terminus), Phospho (STY); **TMT:** TMT 6-plex (K, N-Terminal), Oxidation (M), Phospho (STY)	Protein database	Contaminants-23042018
Static modification	Carboamidomethyl (C)	Results	
Minimum peptide length	6	Protein	Master
Percolator		Protein confidence	High
Target FDR (Strict)	0.01	Unique protein	Depends on data
Target FDR (Relaxed)	0.05		

approaches also depends on the experimental designs and type of data sets.

Targeted Proteomics Data Analysis

The steps for data analysis of MRM and PRM are the same. We have performed data analysis using Skyline (Skyline-daily 19.1.9.350).

i. Open the skyline document containing the list of transitions.
ii. Now click on import results located under the file tab as shown in **Figure S9A**.
iii. Locate the folder containing the results and upload the files at once. You would see a window like the one shown in **Figure S9B**.
iv. Once the import is completed, look at the retention times of the peaks that Skyline detects automatically. To ensure that the right peak has been detected, go to the "View" tab and select replicate comparison under the retention time option.
v. Now correct the retention times of peptides that have been wrongly annotated by Skyline.

CRITICAL STEP: Consider the dot *p* values, shape and intensity of the peak among the many other parameters while deciding on the right peak. The re-annotation involves dragging the mouse cursor below the X-axis from the start time to the end time of the eluted peak.

vi. Once the re-annotation is complete and the areas of all the peaks have been corrected, save the document.
vii. Export the data and perform statistical data analysis.

TIMING

Steps 1-2, Plasma sample preparation: 20 min
Steps 3-9, Depletion of high abundant proteins: 1 h
Steps 10-14, Protein quantification & sample preparation for digestion: 2 h
Steps 15-17, Enzymatic digestion of plasma proteins: 6-8 h
Steps 18-27, Desalting of digested peptides: 1 h per 5 samples
Steps 28-38, Quantification of desalted peptides: 10 min
Steps 39A-39B, Experiment 1: Label-free quantitation (LFQ) of plasma sample: 2 h

Steps 40A-40F, Experiment 2: Label-based quantitation (iTRAQ 4-plex/ TMT 6-plex) of plasma sample: 5 h
Steps 41A-41C, Experiment 3: Multiple Reaction Monitoring (MRM) assay: 1.5 h
Steps 42A-42C, Experiment 4: Parallel Reaction Monitoring (PRM) assay: 1.5 h
Steps 43A-43B, Data analysis: around 1 d

Troubleshooting

Steps	Problem	Possible Reason	Solution
Step 9	Incomplete removal of albumin or IgG	Sample exceeds binding capacity	Reduce the amount of sample processed
		Incomplete binding	Increase incubation time
		The sample is not mixed properly during incubation	Mix the sample with resin by gentle end-over-end mixing and make sure that the sample is mixing with the resin during the incubation period
Step 15–17	Incomplete digestion of proteins	pH is not adjusted to ∼8	Adjust the pH of the sample by adding 100 mM ammonium bicarbonate.
		Urea concentration can be more than >1 M	Dilute the sample by adding 100 mM ammonium bicarbonate.
Step 18	Too much background	Buffer, salt interference	Clean-up sample with desalting column (Empore C18 disk, Pierce C18, Spin columns)
		Keratin contamination	Use fresh buffer and always wear gloves
Step 40	Less no. of proteins identified in case of iTRAQ/TMT	Less amount of peptide considered for labeling	Label at least 50 μg of each iTRAQ/TMT reagent
		Labeling efficiency is not good	Check the pH before adding iTRAQ/TMT reagents and incubate it 2 h with intermittent vortexing after every 10 min
Steps 40	Chromatogram is not good	LC setting	Optimize LC setting for your sample
		Salt and other contaminants	Clean-up sample with desalting column (Empore C18 disk, Pierce C18, Spin columns)
		Column issue	Replace the column
Step 40	Fragmentation of the peptides is not good	Poor ionization	Clean the front end of the mass spectrometer
		Bubble issue at the tip of column	Increase the voltage till 2.3 kV, if not resolved, change the analytical column

RESULTS

One of the major challenges of cancer plasma proteomics has been its inability to discover markers with clinical implications. However, improvement in instrumentation and mass spectrometry-based platforms have contributed to the revival of plasma proteomics (36–39). Currently, several proteomics techniques are being used for MS-based quantitation of plasma proteins for different cancer projects. This study provides a complete proteomics workflow for the discovery and validation of potential biomarker candidates from plasma samples using mass spectrometry. Additionally, the study also provides an optimized sample preparation strategy to get decent coverage of the plasma proteome, which is essential for cancer biomarker discovery projects.

We used a 120 min LC gradient for label-free quantitation and (**Figure S10A**) detected 2332 peptides corresponding to 241 proteins with at least one unique peptide at 1% FDR (**Table S1**). We identified 183 proteins common in all three samples (**Figure 2A**). The heatmap and correlation matrix indicate high levels of consistency (Pearson r value $>$ 0.99) (**Figure 2B** and **Figure S10B**) between the technical replicates (R1, R2, and R3) of different biological samples (Sample A, B, and C). We observed an average of 965 peptides and 170 proteins below than 20% coefficient of variation (CV) (**Figures 2C,D**). In case of iTRAQ experiment, we have labeled varying amounts of peptides (100, 50, 25, 12.5 μg) using iTRAQ reagents to check the minimum amount of peptide to be labeled and observed minimum 50 μg amount of peptide could be used for the good quantitative proteomics experiment (**Figure S11**). However, the number of proteins identified in 114 labeled samples was relatively lower than the other three labels, i.e., 115, 116, and 117. This could be a result of labeling a significantly low number of peptides with the 114-label compared to the other three labels. Around 219 proteins were identified and quantified using iTRAQ-based multiplexed quantitative proteomics (**Figure 3A**, **Table S2**). In TMT experiments, we identified 376 proteins, and 182 proteins were common across all the three quantitative proteomics techniques (LFQ, iTRAQ 4-plex, and TMT 6-plex) (**Figure 3B**, **Table S3**). Studies performing in-depth comparisons of label-free and label-based quantitation (37, 40–42) are also available. We observed a slight increase in the identification of proteins using fractionated samples (six fractions) of TMT 6-plex experiment in comparison to label-free quantitation and iTRAQ 4-plex with a 43.3% overlap between the proteins identified using all three approaches (**Figure 3B**). Further, LFQ provides the flexibility of analyzing clinical samples processing and running as or when available and generating individual datasets. Obtained peptides/protein datasets could be analyzed in different contexts based on IHC, radiology, and other known clinical parameters to address various clinical questions in cancer biology.

The recent developments in the field of targeted proteomics are showing promises in bridging the gap between biomarker discovery and validation of the potential biomarkers (15, 30, 43). We have provided here a workflow for targeted proteomics using

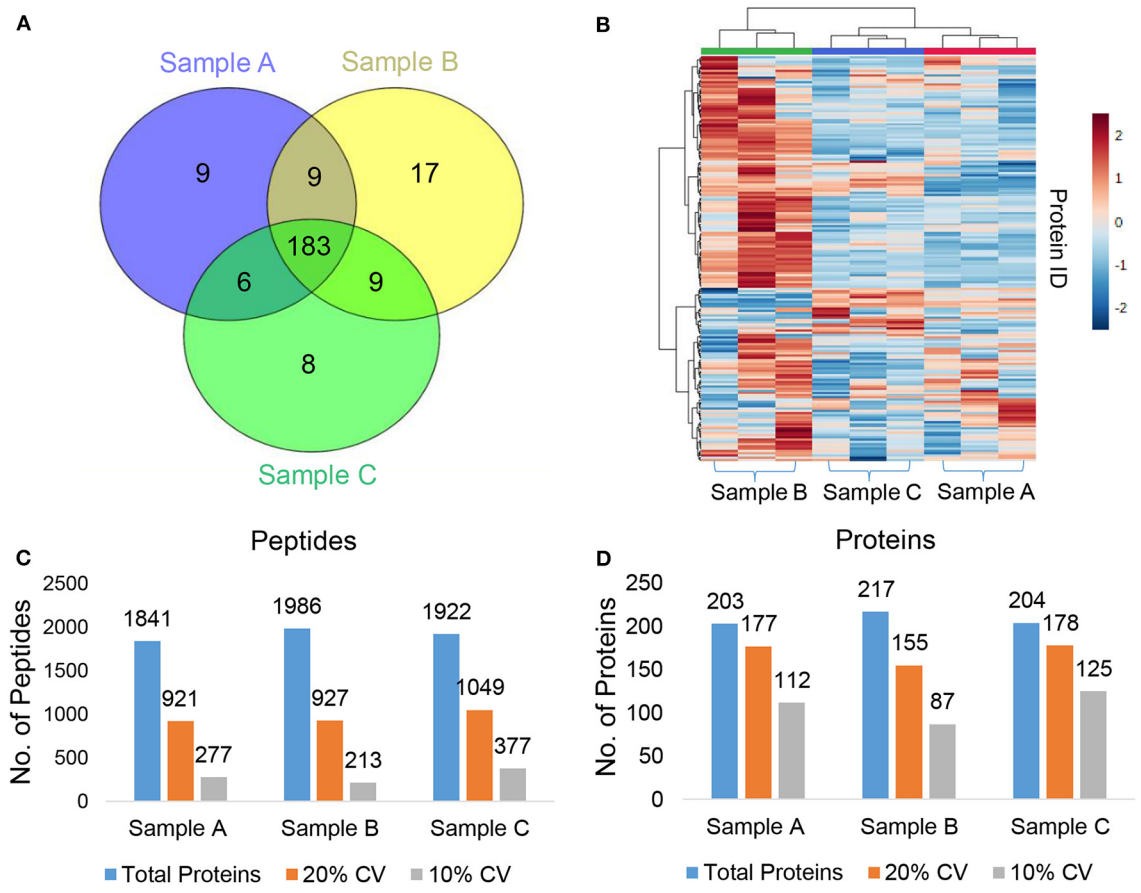


FIGURE 2 | Label-free quantitation (LFQ) of plasma proteins. **(A)** Venn diagram representing the common and unique proteins across the different plasma samples. **(B)** Heatmap showing abundances of the identified proteins in each technical replicate of the three pooled plasma samples (Sample A, B, and C). **(C,D)** The total number of identified **(C)** peptides **(D)** proteins, no. of proteins below 20%, and 10% coefficient of variation (CV) at 1% FDR.

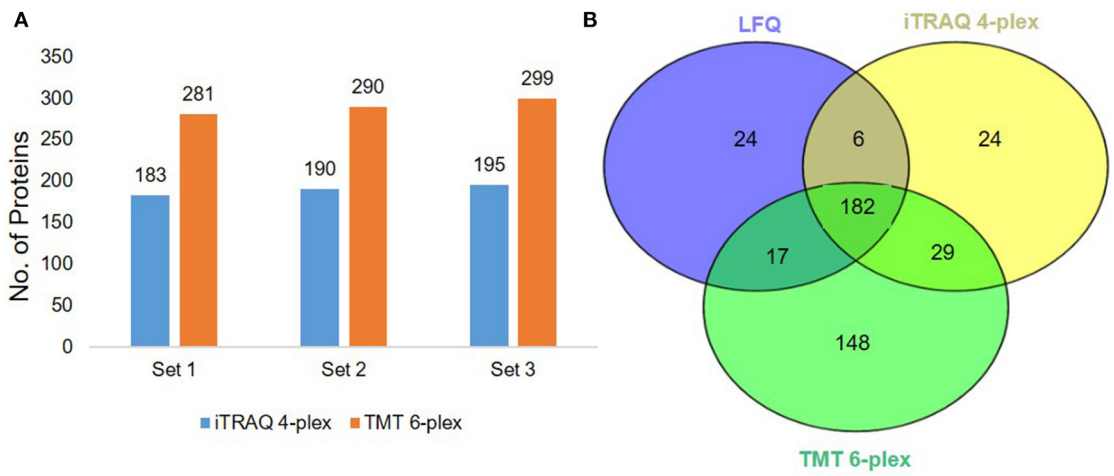


FIGURE 3 | Label-based quantitation of plasma proteins. **(A)** The total no. of identified proteins iTRAQ 4-plex and TMT 6-plex experiments at 1% FDR. **(B)** Venn diagram representing the common and unique proteins across the three different approaches i.e., label-free quantitation, iTRAQ 4-plex, and TMT 6-plex.

PRM and MRM approaches. The main difference between PRM and MRM is that we have to define the transition list in case of MRM and isolation list in case of PRM (**Figure S12**). The abundance of each peptide and reproducibility of retention time across three technical replicates of three biological samples are represented in **Figures 4A,B**. The representative peak intensities,

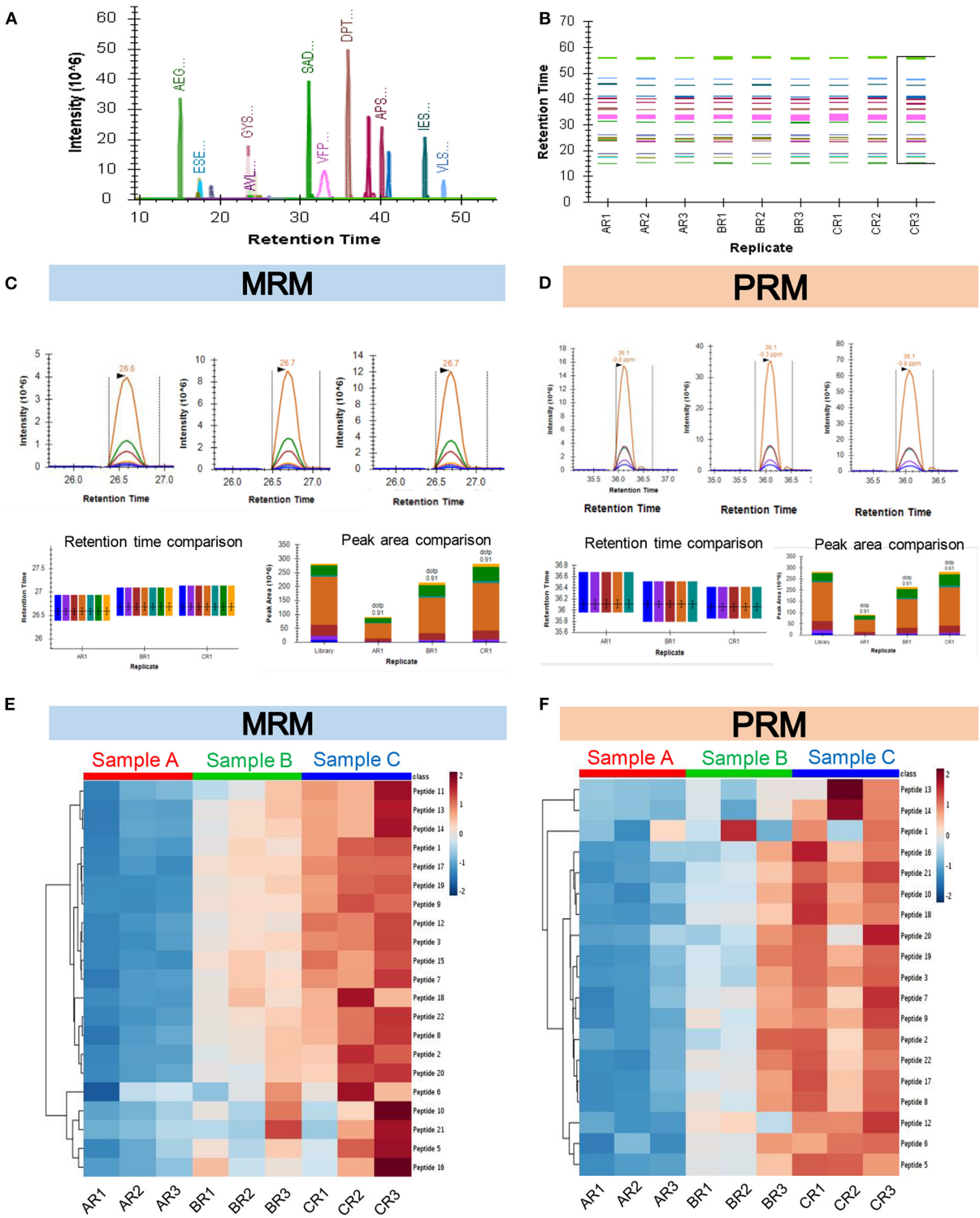


FIGURE 4 | Quantification of synthetic peptides using targeted MS approaches. **(A)** Representative image for all synthetic peptides quantified in the PRM experiment. The plot shows the retention times of the peptides along the x-axis vs. their corresponding intensities along the y-axis. **(B)** Plot showing consistency in retention times of the peptides across all the replicates. **(C)** Representative peak intensities, Retention times, and Peak areas across various dilutions for the Peptide DPTFIPAPIQAK as observed in the MRM experiment. **(D)** Representative peak intensities, Retention times, and Peak areas across various dilutions for the Peptide DPTFIPAPIQAK as observed in the PRM experiment. **(E,F)** The heatmap showing abundances of the identified proteins in each technical replicate (R1, R2, R3) for three biological pools of plasma samples in increasing order of concentration (A, B, and C) as seen in the MRM **(E)** and PRM **(F)** experiments.

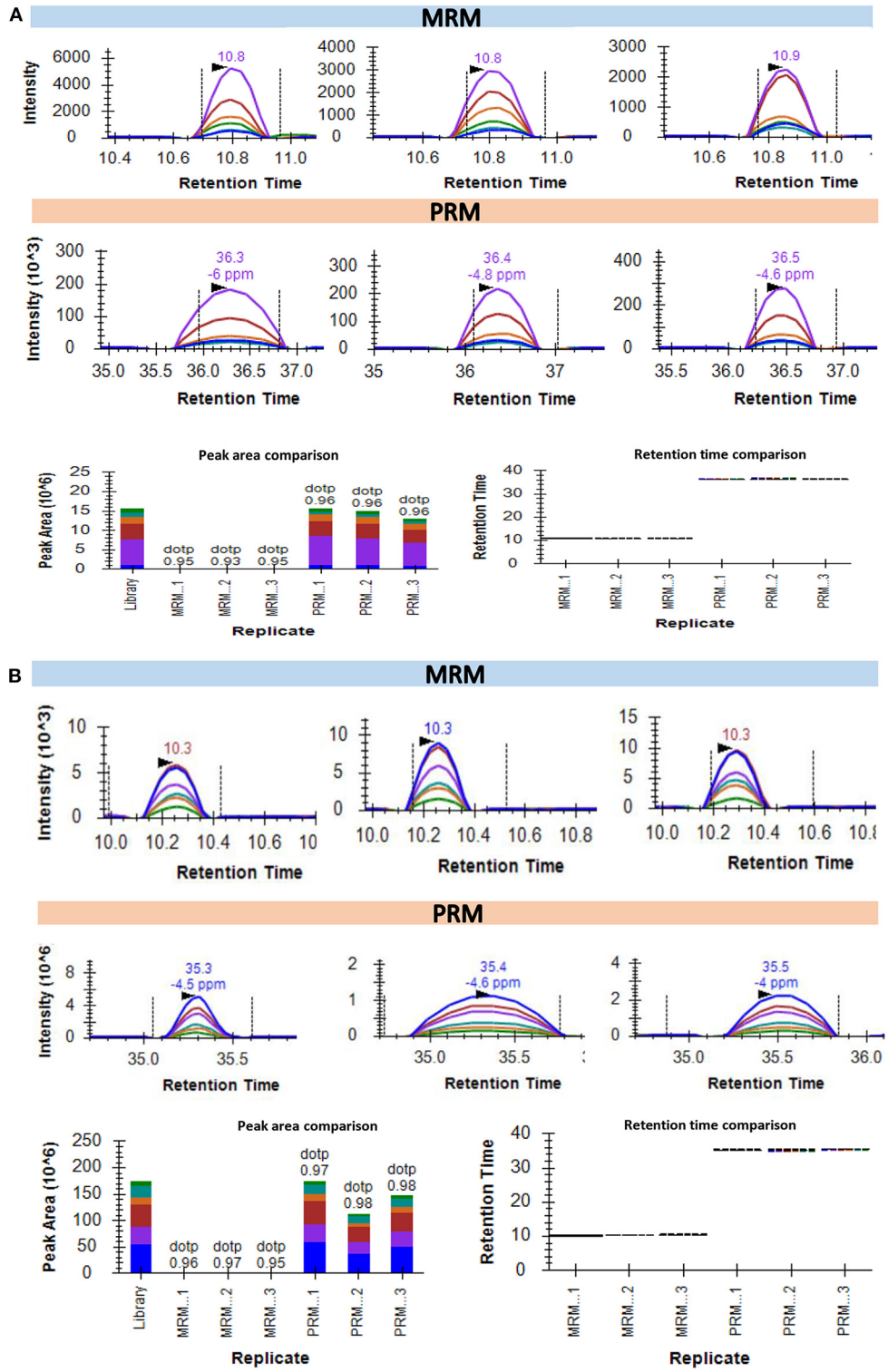


FIGURE 5 | Quantification of the endogenous peptides in plasma samples using targeted MS approaches. **(A)** Intensities of AGALNSNDAFVLK for protein Gelsolin-1 in 3 plasma samples as detected using MRM and PRM. The bottom left part of the panel represents the comparative peak areas of the peptide in each of the 3 samples using both the techniques. The bottom right of the panel shows the consistency of retention times across the biological replicates in both the techniques. **(B)** Intensities of SGLSTGWTQLSK for protein Alpha-1B-glycoprotein in 3 plasma samples as detected using MRM and PRM. The bottom left part of the panel representing the comparative peak areas of the peptide in each of the 3 samples using both the techniques. The bottom right of the panel shows a comparative analysis.

retention times and peak areas across various dilutions for the peptide DPTFIPAPIQAK as observed in the MRM experiment and PRM experiment (**Figures 4C,D**). The intensity of the synthetic peptides using MRM and PRM was found to be correlated with the levels of synthetic peptides spiked into samples A, B, and C (**Figures 4E,F**). We monitored the levels of a few potential cancer biomarkers in plasma samples using MRM, and PRM approaches. The peptide AGALNSNDAFVLK from Gelsolin-1 and SGLSTGWTQLSK from Alpha-1B-glycoprotein showed a good response (Intensities of 10^3 in MRM and 10^6 PRM) and good spectral match with library (dotp value > 0.93) in both the targeted approaches (**Figures 5A,B**).

DISCUSSION

Quantitative approaches involving ultra-sensitive mass spectrometers, which are presented as the pinnacle of promising proteomics technologies, are undoubtedly one of the most widely used approaches in biomarker discovery in recent years. The integrated quantitative proteomics pipeline combining global and targeted approaches described here could be extremely useful in cancer biomarker discovery and validation in plasma samples without a need for any separate immunoassay-based validation method.

Preanalytical variables introduced during blood collection, plasma separation, and storage conditions can adversely influence the quantification of proteins in plasma samples (44), and thereby the outcome of the overall analysis. Potential cancer biomarkers are often very low-abundance proteins and the numbers of detectable proteins are restricted by the complexity of plasma or serum proteome (6, 45, 46). Therefore, it requires extensive depletion of the high-abundance proteins and fractionation methods to obtain comprehensive coverage of the plasma proteome, which certainly introduces substantial experimental time and cost in such quantitative proteomics workflow. In general, the establishment of any clinically relevant protein biomarker panel requires analysis involving large clinical cohorts, including multiple types of control populations (2, 23), which is more crucial for cancer biomarker based projects due to the inter- and intra-tumoral heterogeneity. However, the sample throughput of the discovery phase quantitative proteomics is still moderate and not adequately efficient to satisfy this need (47). Finally, due to the requirement of sophisticated instrumentation and experienced personnel, such MS-based quantitative proteomics workflow is not suitable for routine screening of blood samples in clinical setups.

Analysis of plasma proteome using two complementary quantitation methods as described here provided a satisfactory coverage. Despite advancements in biomarker discovery, there is still no consensus on whether pooling serum samples for shotgun proteomics experiments is always advisable in the discovery phase. While there are many studies that have used serum pooling as a strategy for cancer biomarker discovery (48–51), there also exist studies which advocate otherwise (52, 53). Pooling of clinical samples are often practiced in quantitative proteomics analysis when large numbers of samples need to be studied or there is not an adequate amount of each sample for individual analysis. If sample pooling is performed during the discovery phase of the analysis, it is essential to validate the results in individual diseased and control samples selected randomly from the pooled populations.

In this workflow, the discovery phase experiments [Label-free (LFQ) and Label-based (iTRAQ or TMT)] were performed using an Orbitrap Fusion instrument. The targeted (validation) experiments were performed using two different platforms: multiple reaction monitoring (MRM) using a Triple Quadrupole instrument, and parallel reaction monitoring (PRM) using an Orbitrap Fusion instrument. These two techniques are based on similar principles, and the choice of the method is largely reliant on the type of instrument available to the users. Plasma abundance of a potential cancer biomarker—Alpha-1B-glycoprotein was monitored in the pooled samples and further validated in individual samples using MS-based targeted approaches (**Figure S13**). Using this integrated quantitative proteomics workflow we were able to quantify the relatively low abundant plasma proteins as well (**Figure S14**). The targeted approaches were found to be much superior in terms of quantification accuracy in comparison to the shotgun proteomics approaches. While MRM experiments can be carried out on low-resolution instruments like the triple quadrupole LC-MS (QqQ LC-MS), PRM experiments require the use of high-resolution LC-MS instruments with the QTOF or Q-Orbitrap configuration. Taken together, we conclude that plasma proteomics-based cancer biomarker projects could heavily benefit from detailed workflows of quantitative and targeted proteomics provided in this study. We have demonstrated here multiple possible quantitative approaches in the discovery and validation phases of this combined workflow, but all the methods are not required to be performed simultaneously. Different combinations including any of these discovery and validation phase approaches, could be implemented in biomarker research. Selection of the specific label-based or label-free quantification approach for discovery workflow and MRM or PRM for targeted workflow may depend on the key biological question to be addressed, number of samples, and availability of MS instruments and resources.

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.ebi.ac.uk/pride/archive/, PXD017834, http://www.peptideatlas.org/ (54), PASS01619.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by institutional review boards and ethics committee of the Indian Institute of Technology Bombay

(IITB-IEC/2016/026). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VK, SS, and SR conceived and designed the experiments. VK performed the MS-based quantitative proteomics experiments and data were analyzed by VK, SR, and SG. The manuscript was written by VK, SR, SG, and SS. All authors agreed on the interpretation of data and approved the final version of the manuscript.

FUNDING

This work was supported by Uchhatar Avishkar Yojana (UAY-MHRD), project #34_IITB (2016) to SS. VK was supported by the IIT Bombay fellowship.

ACKNOWLEDGMENTS

We are grateful to Nikita Gahoi and Shalini Aggarwal from the Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay for their insights and suggestions regarding the quantitative and targeted proteomics experiments. We acknowledge the MASSFIITB Facility at IIT Bombay supported by the Department of Biotechnology (BT/PR13114/INF/22/206/2015) to carry out all MS-related experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.543997/full#supplementary-material

REFERENCES

1. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics.* (2002) 1:845–67. doi: 10.1074/mcp.r200007-mcp200
2. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol.* (2017) 13:942. doi: 10.15252/msb.20156297
3. Jin X, Lin M, Zhang H, Han Y, He Y, Zhang Q, et al. Serum biomarkers of colorectal cancer with AU and NP20 chips including a diagnosis model. *Hepatogastroenterology.* (2012) 59:124–129. doi: 10.5754/hge11287
4. Taguchi A, Hanash SM. Unleashing the power of proteomics to develop blood-based cancer markers. *Clin Chem.* (2013) 59:119–26. doi: 10.1373/clinchem.2012.184572
5. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. *Nature.* (2008) 452:571–9. doi: 10.1038/nature06916
6. Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S. Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. *Proteomics.* (2011) 11:2139–61. doi: 10.1002/pmic.201000460
7. Song X, Bandow J, Sherman J, Baker JD, Brown PW, McDowell MT, et al. iTRAQ experimental design for plasma biomarker discovery. *J Proteome Res.* (2008) 7:2952–8. doi: 10.1021/pr800072x
8. Keshishian H, Burgess MW, Specht H, Wallace L, Clauser KR, Gillette MA, et al. Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry. *Nat Protoc.* (2017) 12:1683–701. doi: 10.1038/nprot.2017.054
9. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods.* (2012) 9:555–66. doi: 10.1038/nmeth.2015
10. Lin D, Alborn WE, Slebos RJC, Liebler DC. Comparison of protein immunoprecipitation-multiple reaction monitoring with ELISA for assay of biomarker candidates in plasma. *J Proteome Res.* (2013) 12:5996–6003. doi: 10.1021/pr400877e
11. Peti APF, Locachevic GA, Prado MKB, de Moraes LAB, Faccioli LH. High-resolution multiple reaction monitoring method for quantification of steroidal hormones in plasma. *J Mass Spectrom.* (2018) 53:423–31. doi: 10.1002/jms.4075
12. Ronsein GE, Pamir N, von Haller PD, Kim DS, Oda MN, Jarvik GP, et al. Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J Proteomics.* (2015) 113:388–99. doi: 10.1016/j.jprot.2014.10.017
13. Jaffe JD, Keshishian H, Chang B, Addona TA, Gillette MA, Carr SA. Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. *Mol Cell Proteomics.* (2008) 7:1952–62. doi: 10.1074/mcp.M800218-MCP200
14. Hoofnagle AN, Becker JO, Oda MN, Cavigiolio G, Mayer P, Vaisar T. Multiple-reaction monitoring-mass spectrometric assays can accurately measure the relative protein abundance in complex mixtures. *Clin Chem.* (2012) 58:777–81. doi: 10.1373/clinchem.2011.173856
15. Whiteaker JR, Lin C, Kennedy J, Hou L, Trute M, Sokal I, et al. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat Biotechnol.* (2011) 29:625–34. doi: 10.1038/nbt.1900
16. Liebler DC, Zimmerman LJ. Targeted quantitation of proteins by mass spectrometry. *Biochemistry.* (2013) 52:3797–806. doi: 10.1021/bi400110b
17. Aebersold R, Burlingame AL, Bradshaw RA. Western blots vs. selected reaction monitoring assays: time to turn the tables? *Mol Cell Proteomics.* (2013) 12:2381–2. doi: 10.1074/mcp.E113.031658
18. Atak A, Khurana S, Gollapalli K, Reddy PJ, Levy R, Ben-Salmon S, et al. Quantitative mass spectrometry analysis reveals a panel of nine proteins as diagnostic markers for colon adenocarcinomas. *Oncotarget.* (2018) 9:13530–44. doi: 10.18632/oncotarget.24418
19. Jadhav M, Nayak M, Kumar S, Venkatesh A, Patel SK, Kumar V, et al. Clinical proteomics and cytokine profiling for dengue fever disease severity biomarkers. *OMICS.* (2017) 21:665–77. doi: 10.1089/omi.2017.0135
20. Ray S, Kumar V, Bhave A, Singh V, Gogtay NJ, Thatte UM, et al. Proteomic analysis of *Plasmodium falciparum* induced alterations in humans from different endemic regions of India to decipher malaria pathogenesis and identify surrogate markers of severity. *J Proteomics.* (2015) 127:103–13. doi: 10.1016/j.jprot.2015.04.032
21. Ray S, Patel SK, Venkatesh A, Bhave A, Kumar V, Singh V, et al. Clinicopathological analysis and multipronged quantitative proteomics reveal oxidative stress and cytoskeletal proteins as possible markers for severe vivax malaria. *Sci Rep.* (2016) 6:24557. doi: 10.1038/srep24557
22. Ray S, Patel SK, Venkatesh A, Chatterjee G, Ansari NN, Gogtay NJ, et al. Quantitative proteomics analysis of *Plasmodium vivax* induced alterations in human serum during the acute and convalescent phases of infection. *Sci Rep.* (2017) 7:4400. doi: 10.1038/s41598-017-04447-5
23. Sharma S, Ray S, Moiyadi A, Sridhar E, Srivastava S. Quantitative proteomic analysis of meningiomas for the identification of surrogate protein markers. *Sci Rep.* (2014) 4:7140. doi: 10.1038/srep07140
24. Sharma S, Ray S, Mukherjee S, Moiyadi A, Sridhar E, Srivastava S. Multipronged quantitative proteomic analyses indicate modulation of various signal transduction pathways in human meningiomas. *Proteomics.* (2015) 15:394–407. doi: 10.1002/pmic.201400328
25. Mertins P, Tang LC, Krug K, Clark DJ, Gritsenko MA, Chen L, et al. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass

spectrometry. *Nat Protoc.* (2018) 13:1632–61. doi: 10.1038/s41596-018-0006-9
26. O'Connell JD, Paulo JA, O'Brien JJ, Gygi SP. Proteome-wide evaluation of two common protein quantification methods. *J Proteome Res.* (2018) 17:1934–42. doi: 10.1021/acs.jproteome.8b00016
27. Ray S, Lach R, Heesom KJ, Valekunja UK, Encheva V, Snijders AP, et al. Phenotypic proteomic profiling identifies a landscape of targets for circadian clock-modulating compounds. *Life Sci Alliance.* (2019) 2:603. doi: 10.26508/lsa.201900603
28. Reddy PJ, Ray S, Sathe GJ, Gajbhiye A, Prasad TSK, Rapole S, et al. A comprehensive proteomic analysis of totarol induced alterations in Bacillus subtilis by multipronged quantitative proteomics. *J Proteomics.* (2015) 114:247–62. doi: 10.1016/j.jprot.2014.10.025
29. Martinez-Garcia E, Lesur A, Devis L, Cabrera S, Matias-Guiu X, Hirschfeld M, et al. Targeted proteomics identifies proteomic signatures in liquid biopsies of the endometrium to diagnose endometrial cancer and assist in the prediction of the optimal surgical treatment. *Clin Cancer Res.* (2017) 23:6458–67. doi: 10.1158/1078-0432.CCR-17-0474
30. Keshishian H, Burgess MW, Gillette MA, Mertins P, Clauser KR, Mani DR, et al. Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol Cell Proteomics.* (2015) 14:2375–93. doi: 10.1074/mcp.M114.046813
31. Venkatesh A, Lahiri A, Reddy PJ, Shastri J, Bankar S, Patankar S, et al. Identification of highly expressed *Plasmodium vivax* proteins from clinical isolates using proteomics. *Prot Clin Appl.* (2018) 12:1700046. doi: 10.1002/prca.201700046
32. Awasthi G, Tyagi S, Kumar V, Patel SK, Rojh D, Sakrappanavar V, et al. A proteogenomic analysis of haptoglobin in malaria. *Proteomics Clin Appl.* (2018) 12:e1700077. doi: 10.1002/prca.201700077
33. Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian W-J, Webb-Robertson B-JM, et al. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J Proteome Res.* (2006) 5:277–86. doi: 10.1021/pr050300l
34. O'Rourke MB, Town SEL, Dalla PV, Bicknell F, Koh Belic N, Violi JP, et al. What is normalization? The strategies employed in top-down and bottom-up proteome analysis workflows. *Proteomes.* (2019) 7:19. doi: 10.3390/proteomes7030029
35. Välikangas T, Suomi T, Elo LL. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief Bioinform.* (2018) 19:1–11. doi: 10.1093/bib/bbw095
36. Beer LA, Liu P, Ky B, Barnhart KT, Speicher DW. Efficient quantitative comparisons of plasma proteomes using label-free analysis with MaxQuant. *Methods Mol Biol.* (2017) 1619:339–52. doi: 10.1007/978-1-4939-7057-5_23
37. Patel VJ, Thalassinos K, Slade SE, Connolly JB, Crombie A, Murrell JC, et al. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res.* (2009) 8:3752–9. doi: 10.1021/pr900080y
38. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics.* (2004) 3:1154–69. doi: 10.1074/mcp.M400129-MCP200
39. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem.* (2003) 75:1895–904. doi: 10.1021/ac0262560
40. Li Z, Adams RM, Chourey K, Hurst GB, Hettich RL, Pan C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. *J Proteome Res.* (2012) 11:1582–90. doi: 10.1021/pr200748h
41. Megger DA, Pott LL, Ahrens M, Padden J, Bracht T, Kuhlmann K, et al. Comparison of label-free and label-based strategies for proteome analysis of hepatoma cell lines. *Biochim Biophys Acta.* (2014) 1844:967–76. doi: 10.1016/j.bbapap.2013.07.017
42. Sandberg A, Branca RMM, Lehtiö J, Forshed J. Quantitative accuracy in mass spectrometry based proteomics of complex samples: the impact of labeling and precursor interference. *J Proteomics.* (2014) 96:133–44. doi: 10.1016/j.jprot.2013.10.035
43. Makawita S, Diamandis EP. The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: current strategies for candidate verification. *Clin Chem.* (2010) 56:212–22. doi: 10.1373/clinchem.2009.127019
44. Rai AJ, Vitzthum F. Effects of preanalytical variables on peptide and protein measurements in human serum and plasma: implications for clinical proteomics. *Expert Rev Proteomics.* (2006) 3:409–26. doi: 10.1586/14789450.3.4.409
45. Ebert MPA, Korc M, Malfertheiner P, Röcken C. Advances, challenges, and limitations in serum-proteome-based cancer diagnosis. *J Proteome Res.* (2006) 5:19–25. doi: 10.1021/pr050271e
46. Issaq HJ, Xiao Z, Veenstra TD. Serum and plasma proteomics. *Chem Rev.* (2007) 107:3601–20. doi: 10.1021/cr068287r
47. Qian W-J, Jacobs JM, Liu T, Camp DG, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics.* (2006) 5:1727–44. doi: 10.1074/mcp.M600162-MCP200
48. Zhong M-E, Chen Y, Xiao Y, Xu L, Zhang G, Lu J, et al. Serum extracellular vesicles contain SPARC and LRG1 as biomarkers of colon cancer and differ by tumour primary location. *EBioMedicine.* (2019) 50:211–23. doi: 10.1016/j.ebiom.2019.11.003
49. Lobo MDP, Moreno FBMB, Souza GHMF, Verde SMML, Moreira R de A, Monteiro-Moreira AC de O. Label-free proteome analysis of plasma from patients with breast cancer: stage-specific protein expression. *Front Oncol.* (2017) 7:14. doi: 10.3389/fonc.2017.00014
50. Cheng Y, Liu C, Zhang N, Wang S, Zhang Z. Proteomics analysis for finding serum markers of ovarian cancer. *Biomed Res Int.* (2014) 2014:179040. doi: 10.1155/2014/179040
51. Shiromizu T, Kume H, Ishida M, Adachi J, Kano M, Matsubara H, et al. Quantitation of putative colorectal cancer biomarker candidates in serum extracellular vesicles by targeted proteomics. *Sci Rep.* (2017) 7:12782. doi: 10.1038/s41598-017-13092-x
52. Molinari N, Roche S, Peoc'h K, Tiers L, Séveno M, Hirtz C, et al. Sample pooling and inflammation linked to the false selection of biomarkers for neurodegenerative diseases in top-down proteomics: a pilot study. *Front Mol Neurosci.* (2018) 11:477. doi: 10.3389/fnmol.2018.00477
53. Kendziorski C, Irizarry RA, Chen K-S, Haag JD, Gould MN. On the utility of pooling biological samples in microarray experiments. *Proc Natl Acad Sci USA.* (2005) 102:4252–4257. doi: 10.1073/pnas.0500607102
54. Farrah T, Deutsch EW, Kreisberg R, Sun Z, Campbell DS, Mendoza L, et al. PASSEL: the PeptideAtlas SRMexperiment library. *Proteomics.* (2012) 12:1170–5. doi: 10.1002/pmic.201100515

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 © 2020 Kumar, Ray, Ghantasala and Srivastava. 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 Oncology

ORIGINAL RESEARCH
published: 25 September 2020
doi: 10.3389/fonc.2020.01767

Ovarian Blood Sampling Identifies Junction Plakoglobin as a Novel Biomarker of Early Ovarian Cancer

Florian Weiland[1,2,3], Noor A. Lokman[4], Manuela Klingler-Hoffmann[5], Thomas Jobling[6], Andrew N. Stephens[7,8], Karin Sundfeldt[9], Peter Hoffmann[5] and Martin K. Oehler[4,5,10]**

[1] Adelaide Proteomics Centre, The University of Adelaide, Adelaide, SA, Australia, [2] Institute for Photonics and Advanced Sensing, The University of Adelaide, Adelaide, SA, Australia, [3] Department of Microbial and Molecular Systems (M^2S), Laboratory of Enzyme, Fermentation and Brewing Technology (EFBT), KU Leuven, Leuven, Belgium, [4] Discipline of Obstetrics and Gynecology, Adelaide Medical School, Robinson Research Institute, The University of Adelaide, Adelaide, SA, Australia, [5] Future Industries Institute, University of South Australia, Adelaide, SA, Australia, [6] Department of Gynecological Oncology, Monash Medical Centre, Clayton, VIC, Australia, [7] Centre for Cancer Research, Hudson Institute of Medical Research, Clayton, VIC, Australia, [8] Department of Molecular and Translational Sciences, Monash University, Clayton, VIC, Australia, [9] Department of Obstetrics and Gynecology, Sahlgrenska Cancer Center, Institute of Clinical Science, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [10] Department of Gynecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

OPEN ACCESS

Edited by:
Harsha Gowda,
The University of Queensland, Australia

Reviewed by:
Hitoshi Tsuda,
National Defense Medical College, Japan
Suresh Mathivanan,
La Trobe University, Australia

***Correspondence:**
Peter Hoffmann
peter.hoffmann@unisa.edu.au
Martin K. Oehler
martin.oehler@adelaide.edu.au

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 12 April 2020
Accepted: 06 August 2020
Published: 25 September 2020

Citation:
Weiland F, Lokman NA, Klingler-Hoffmann M, Jobling T, Stephens AN, Sundfeldt K, Hoffmann P and Oehler MK (2020) Ovarian Blood Sampling Identifies Junction Plakoglobin as a Novel Biomarker of Early Ovarian Cancer. Front. Oncol. 10:1767. doi: 10.3389/fonc.2020.01767

Ovarian cancer is the most lethal gynecologic malignancy. Early detection would improve survival, but an effective diagnostic test does not exist. Novel biomarkers for early ovarian cancer diagnosis are therefore warranted. We performed intraoperative blood sampling from ovarian veins of stage I epithelial ovarian carcinomas and analyzed the serum proteome. Junction plakoglobin (JUP) was found to be elevated in venous blood from ovaries with malignancies when compared to those with benign disease. Peripheral plasma JUP levels were validated by ELISA in a multicenter international patient cohort. JUP was significantly increased in FIGO serous stage IA+B (1.97-fold increase; $p < 0.001$; $n = 20$), serous stage I (2.09-fold increase; $p < 0.0001$; $n = 40$), serous stage II (1.81-fold increase, $p < 0.001$, $n = 23$) and serous stage III ovarian carcinomas (1.98-fold increase; $p < 0.0001$; $n = 34$) vs. normal controls ($n = 109$). JUP plasma levels were not increased in early stage breast cancer ($p = 0.122$; $n = 12$). In serous ovarian cancer patients, JUP had a sensitivity of 85% in stage IA+B and 60% in stage IA-C, with specificities of 76 and 94%, respectively. A logistic regression model of JUP and Cancer Antigen 125 (CA125) revealed a sensitivity of 70% for stage IA+B and 75% for stage IA-C serous carcinomas at 100% specificity. Our novel ovarian blood sampling – proteomics approach identified JUP as a promising new biomarker for epithelial ovarian cancer, which in combination with CA125 might fulfill the test criteria for ovarian cancer screening.

Keywords: junction plakoglobin, ovarian cancer, early diagnosis, biomarker, proteomics, abundance protein depletion, saturation labeling 2D-DIGE, mass spectrometry

INTRODUCTION

Ovarian cancer (OC) accounts for an estimated 239,000 new cases and 152,000 deaths worldwide each year (1). The high mortality rate of ovarian cancer arises due to the asymptomatic progression of the disease, resulting in over 70% of cases being diagnosed at advanced stage [International Federation of Gynecology and Obstetrics (FIGO) stage III and IV] when the cancer has spread

to the abdominal cavity or to other organs. Detection of OC at an early stage, i.e., when it is still confined to the ovary (FIGO stage I), is associated with a 5-year survival rate of over 90% compared to <30% for women presenting with advanced disease (2, 3). Therefore, the detection of OC at an early stage is the most effective way to improve survival. However, at present no clinically applicable early detection test is available, and population screening is therefore not possible.

CA125, the current gold standard protein biomarker in OC, is only clinically approved to distinguish benign from malignant ovarian lesions and to assess tumor burden. Elevated levels of CA125, however, are only observed in <50% of early stage patients and CA125 is therefore not a useful tumor marker for early ovarian cancer detection (4).

Significant efforts have been undertaken to develop a biomarker-based early detection test for many years. However, a major "quantitative" challenge of OC biomarker research is to identify the cancer at an early stage when it is small and only secretes a very small amount of cancer-specific biomarker into the blood stream. Another significant "qualitative" problem of biomarker detection is that highly abundant proteins tend to obscure the detection of potential biomarkers that are usually in lower concentrations in biofluids such as serum or plasma (5).

To overcome the challenges of OC biomarker detection, we developed a novel approach of combining: (a) Ovarian blood sampling, to obtain blood with higher biomarker concentration downstream of the cancer, with (b) Abundance protein depletion and saturation labeling 2D-DIGE, to identify less abundant proteins in plasma and serum. Our strategy resulted in the identification of a promising new marker for epithelial ovarian cancer, junction plakoglobin (JUP, plakoglobin, γ-catenin).

JUP is a member of the Armadillo family of proteins and responsible for the linking of classical cadherins to α-catenin (6–9). It interacts with the desmosomal cadherins desmoglein-1 and desmocollin (10), anchoring cell-cell adhesion receptors into the cytoskeleton. JUP has never been described as diagnostic biomarker for OC or any other malignancy.

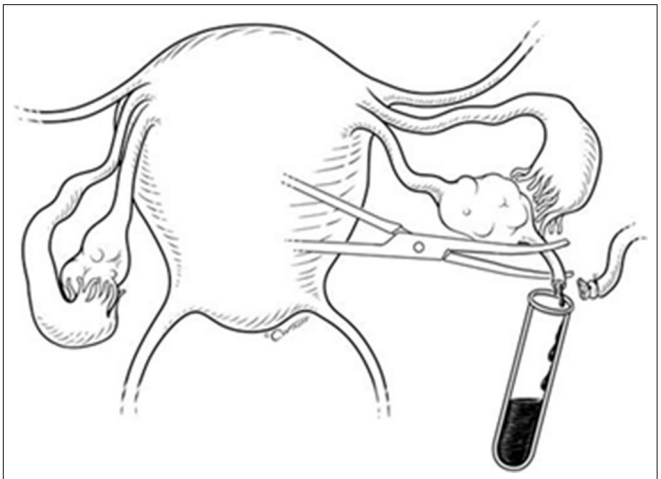


FIGURE 1 | Ovarian blood sampling technique during total abdominal hysterectomy and bilateral salpingo-oophorectomy. Reproduced from Fogle RH, Stanczyk FZ, Zhang X, Paulson RJ. Ovarian androgen production in postmenopausal women. J Clin Endocrinol Metab. 2007; 92:3040–3, with permission of the Endocrine Society.

MATERIALS AND METHODS

Patient Samples

Blood from ovarian and peripheral (cubital) veins was collected intraoperatively with patient consent and approval by the Research Ethics Committee at the Royal Adelaide Hospital, Adelaide, South Australia (Protocol #080304) (**Figure 1**). Ovarian serum samples from 6 patients with stage IA high grade ovarian cancer (2 serous papillary and 4 endometrioid adenocarcinomas), 2 patients with functional ovarian cysts and 4 patients with benign serous cystadenomas were used for biomarker discovery. Ovarian and patient matched peripheral blood was collected into clotting tubes (Greiner Bio-one, Austria), serum EDTA prepared by centrifugation at 3,000 rpm for 15 min at room temperature and supernatant stored at −80°C.

For the confirmation and validation phases, peripheral plasma samples of patients with high grade serous ovarian carcinomas and normal controls were sourced from the Robinson Research Institute, The University of Adelaide, Adelaide, South Australia; Hudson Institute of Medical Research, Clayton, Victoria, Australia; Department of Obstetrics and Gynecology, Sahlgrenska Cancer Center, University of Gothenburg, Gothenburg, Sweden; Ontario Tumor Bank, Toronto, Ontario, Canada, and Precision Med Inc., Solana Beach, California, USA. Samples for the ovarian cancer subtypes cohort (endometrioid, mucinous and clear cell carcinomas) were obtained from the Robinson Research Institute, The University of Adelaide, Adelaide, South Australia (**Table 1**).

Plasma EDTA samples of early stage breast cancer patients were acquired from Conversant Biosciences Inc., Huntsville, Alabama, USA (**Table 1**).

Depletion of Top 14 Abundant Serum Proteins

The 14 most abundant serum proteins were depleted using the multiple affinity removal system (MARS) liquid chromatography (LC) column (Cat. No. 5188-6558, Agilent, Santa Clara, California, USA) according to the manufacturer's recommendations. In brief, 100 µl serum were mixed with 297 µl buffer A (Cat. No. 5185-5987, Agilent) and 3 µl protease inhibitor cocktail (Cat. No. P9599, Sigma-Aldrich, St. Louis, Missouri, USA), then filtered through a 0.22 µm spin filter (Cat. No. CLS8160, Corning, New York, USA) using a centrifuge at 16,000 x g at room temperature for 1 min. Depletion of 160 µl of the filtrate was carried out on an Agilent 1,100 high performance LC (HPLC) using a 4.6 × 100 mm MARS human 14 depletion column. Low abundant protein fraction was collected and mixed 1:5 with 100% (v/v) acetone (ice-cold) and stored at −20°C until further use. Proteins were pelleted by centrifugation for 45 min at 12,000 × g and −9°C. The protein pellet was washed in 3 mL ice-cold 100% (v/v) acetone, stored at −20°C for 1 h then

TABLE 1 | Descriptive statistics of patient samples used for the JUP ELISAs.

Group	FIGO stage	No.	Age	CA125 [kU/L]
A) MULTICENTRE DATASET - PLASMA EDTA (FIGURE 3)				
High grade serous	IA	18	62 (43–84)	54 (6–11,710)
	IB	2	60 (52–68)	1,247 (730–1,763)
	IC	20	60 (45–86)	232 (17–2,323)
	II	23	60 (41–78)	285 (10–2,446)
	III	34	66 (39–84)	816 (47–11,200)
Normal control	–	109	47 (28–86)	9 (2–54)
B) OVARIAN CANCER SUBTYPES - SERUM (FIGURE 4)				
Endometrioid	IA	10	58 (41–68)	50 (12–1,564)
	IC	5	55 (33–90)	141 (113–1,856)
Mucinous	IA	9	61 (37–72)	25 (12–1,570)
	IC	3	45 (29–62)	98 (35–126)
Clear cell	IA	9	61 (43–82)	33 (8–593)
	IB	1	52	20
	IC	1	63	20
Normal control	–	39	46 (26–79)	10 (3–63)
C) BREAST CANCER DATASET - PLASMA EDTA				
Breast cancer	I	6	52 (41–66)	–
	II	6	61 (46–68)	–
High grade serous	IC	3	50 (45–86)	292 (289–390)
	II	2	58 (53–63)	210 (121–298)
Normal control	–	7	60 (37–72)	13 (5–24)

Age and CA125 concentrations are given as median with range in brackets. High grade serous stage IC contains one patient without a known CA125 plasma concentration.

centrifuged for 45 min at 12,000 × g and −9°C. Supernatant was discarded and the protein pellet was air-dried for 10 min to remove residual acetone. Proteins were suspended in 400 µl TUC4% [7 M urea (Cat. No. 1.08487.0500, Merck, Darmstadt, Germany), 2 M thiourea (Cat. No. RPN6301, GE Healthcare, Little Chalfont, UK), 4% (w/v) CHAPS (Cat. No. 10810118001, Roche, Basel, Switzerland), 30 mM Tris (Cat. No. BIO3094T, Astral Scientific, Taren Point, Australia), 1% (v/v) protease inhibitor cocktail (Sigma), 1.1% (v/v) Pefabloc® SC protector reagent (Cat. No. 11873628001, Roche), pH 7.5]. The suspended protein samples were desalted via a 10 kDa cut-off spin filter (Cat. No. VN01H02, Vivacon 500, Sartorius, Göttingen, Germany) by centrifugation for 30 min at 14,000 × g, 15°C. The filter was washed five times by adding 400 µl of TUC4% and centrifugation for 30 min at 14,000 × g, 15°C. The protein concentration of resulting protein sample was measured using an EZQ™ protein quantification kit (Cat. No. R33200, Life Technologies, Carlsbad, California, USA) according to the manufacturer's manual.

Saturation Labeling Two-Dimensional Difference Gel Electrophoresis (2D DIGE)

Amount of protein labels S-200 and S-300 [Cat. No. PR33, NH DyeAgnostics GmbH, Halle (Saale), Germany] and reduction agent Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (NH DyeAgnostics) was optimized according to the manufacturer's recommendation and protocols (11) with 1.5 µl of the TCEP (resuspended in 400 µl of H_2O) and 3 µl of the respective label (S-200 and S-300) chosen for the subsequent saturation DIGE experiment. Six serum samples from the venous flow of malignant stage I epithelial ovarian tumors, the 6 corresponding peripheral serum samples, 6 serum samples from the venous flow of benign ovarian tumors and one internal pooled standard (IPS) were labeled. For this, 5 µg of total proteins from the depleted serum collected from the ovarian venous flow and matched peripheral serum were labeled with S-200. The IPS, consisting of an equal mixture of 5.9 µg of total proteins from each sample was labeled with S-300. Volume of each protein sample was equalized to 9 µl using TUC4%, then proteins were reduced for 1 h at 35°C in the dark using 1.5 µl TCEP for the samples. IPS was diluted with TUC4% to a concentration of 0.56 µg/µl, subsequently 31.9 µl TCEP was added and incubated at 35°C for 1 h in the dark. After reduction, the proteins were labeled using 3 µl of S-200, the IPS was labeled using 63.7 µl of S-300 and incubated for 1 h at 35°C in the dark. For each sample, the reaction was quenched by adding 13.5 µl (IPS: 286.7 µl) of TUC4% supplemented with 2% (w/v) Dithiothreitol (Cat. No. 11583786001, Roche) and 4% (v/v) Pharmalyte 3-10 (Cat. No. 17-0456-01, GE Healthcare).

2D PAGE

Eighteen 24 cm immobilized pH gradient (IPG) strips with a pH range of 3-10NL (Cat. No. 1632043, Bio-rad, Hercules, California, USA) were rehydrated in 500 µl of TUC1% (6 M urea, 2 M thiourea, 1.2% (v/v) 2,2 Dithiodiethanol (Cat. No. 380474, Sigma-Aldrich), 0.5% (v/v) Pharmalyte 3-10 (GE Healthcare) overnight at room temperature, then stored at −80°C until further use. 5µg of protein sample mixed with 5 µg IPS was loaded via anodal cup-loading. Isoelectric focusing (IEF) was carried out using an IPGPhor II (GE Healthcare) with following settings: 150 V for 1 h, 300 V for 1 h, 600 V for 1.5 h, increase to 8,000 V by gradient over 2 h, 24.000 Vh at 8,000 V, under exclusion of light. Current was limited to 50 µA per IPG strip. After IEF, IPG strips were stored at −80°C until further use. SDS-PAGE was carried out as described earlier (12).

Image Acquisition

After SDS-PAGE, gels were scanned using a Typhoon Trio Imager (GE Healthcare). S-200 channel was scanned with a photomultiplier tension (PMT) of 600 V, an emission window of 580 nm BP30, excitation using a green laser (532 nm), S-300 was scanned with a PMT of 600 V, an emission window 670 nm BP30 and excitation with a red laser (633 nm).

DIGE Data Analysis

Before spot detection, acquired images were warped using Robust Automated Image Normalization (13, 14) applying the recommended standard settings. All gels were aligned to the IPS (S-300 channel) of Gel01 (**Figure S1**). Protein spot detection was carried out using DeCyder 7.0 (GE Healthcare, RRID: SCR_014592). On average 2,359 (range: 2,174–2,715) spots per gel were detected. The normalized spot volumes were exported using the DeCyder XML toolbox (GE Healthcare), standardized

against the spot volume of the corresponding spot in the IPS (S-300) channel and log10 transformed. Data was analyzed using R (version 3.6.2, The R Foundation for Statistical Computing, RRID: SCR_001905) (15) using the additional libraries plyr (16) and reshape (17).

Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)

Primary candidate spots (736, 1,252, 1,295 and 1,507) were excised from eight gels using an Ettan™ SpotPicker (GE Healthcare) and the respective spots were combined. Proteins were digested using Trypsin (Cat. No. V5111, Promega, Madison, Wisconsin, USA). Tryptic peptides were dried using a SpeedVac (Thermo-Fisher, Waltham, Massachusetts, USA) and suspended in 12 µl of FA2 (2% (v/v) acetonitrile (Cat. No. 1.00029.1000, Merck), 0.1% (v/v) formic acid (Cat. No. 5438040100, Sigma-Aldrich)). LC-MS/MS was carried out using an Impact HD mass spectrometer (Bruker Daltonics, Bremen, Germany) as described earlier (18), deviating in applying a flow-rate of 0.3 µl/min and a 70 min gradient for peptide separation. Ten µl of peptide mixture was injected.

Protein Identification

Acquired mass spectrometry data was converted to Mascot generic format using ProteoWizard version 3.0.20093 (19) and searched against the SwissProt human database (downloaded 03/12/2019; 42,386 entries) using Comet release 2019.01 rev. 4 (20). Precursor mass tolerance was set to 20 ppm, fragment bin tolerance to 0.02. Variable modifications of oxidation of methionine and fixed modification of Cy3 saturation DIGE label on cysteine were specified, with the digestion enzyme specified as trypsin omitting the proline rule with 2 allowed missed cleavages. Peptides with an e-value below 0.05 were regarded as significant, for protein identification only proteins with at least two significant unique peptide identifications were considered. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (21) partner repository with the dataset identifier PXD018417. Mass spectrometry data analysis scripts can be downloaded from https://github.com/medardus333/OvCa_Markers.

Enzyme-Linked Immunosorbent Assay (ELISA)

JUP ELISA kits were purchased from MyBiosource, San Diego, California, USA (Cat. No. MBS2018947). ELISAs were performed according to the manufacturer's protocol with the modification of using 300 µl wash solution (instead of 350 µl). Plate design was randomized using R. Serum or plasma EDTA samples for JUP ELISA were diluted 1:10 in 1 x PBS pH 7.4. ELISA plates were developed for 15 min, afterwards absorbance at 450 nm was measured using a Biotrak II Reader (GE Healthcare). A 4-parameter logistic regression curve was fitted to the absorbance values of the standards for each ELISA plate and used to calculate the JUP concentration of the samples. For the ELISA measuring the JUP concentration in early stage breast cancer, the calibration point with the highest JUP concentration (10 ng/mL) had to be removed from this standard curve as a too low absorbance reading caused the curve fitting to fail. Values between 50 and 100 ng/mL JUP in this ELISA are therefore extrapolated. However, the calibration curves from the earlier conducted JUP ELISA clearly exhibit linear behavior in this range and give precedent to assume a reasonable accuracy of the extrapolated part of the calibration curve used in the early stage breast cancer ELISA (**Figure S2**). In the multicenter ELISAs, inter-assay coefficient of variance (CV) after calibration using their respective normal control median JUP concentrations was calculated as 6.16% and the median intra-assay CV as 3.96%. For all ELISAs, difference in mean between groups was tested using a two-sided *t*-test, correlations were calculated using a Spearman rank correlation. All analysis steps were performed using R with additional usage of the following packages: drc (22) for calibration curve fitting; ggplot2 (23), ggbeeswarm (24), extrafont (25) and wesanderson (26) for data visualization, dplyr (27), matrixStats (28), and pROC (29) for data analysis and Receiver Operator Characteristics (ROC). The sensitivity at 99.6% specificity level is calculated from a linear regression between the two nearest datapoints spanning 99.6% specificity, the same holds true for specificities at 75% sensitivity in case there was no exact 75% sensitivity datapoint. Positive predictive values (PPV) and negative predictive values (NPV) were calculated using a prevalence of 1:2,500 at a fixed sensitivity level of 75% (4).

RESULTS

Saturation 2D DIGE

Our first hypothesis was that proteins which are higher abundant in the venous backflow from early ovarian cancers than in peripheral blood of the same patient are potential markers for early detection. For statistical analysis a paired, one-sided *t*-test was applied. Protein spot volumes exhibiting a $p < 0.05$ and positive fold-change were regarded as significantly higher abundant in ovarian venous backflow. Our second aim was to control for potential increase of protein spot volumes in the ovarian venous backflow due to the surgical procedure. To address this issue, we hypothesized that potential markers for early stage ovarian cancer are higher abundant in venous backflow from malignant tumors than in the venous backflow from benign tumors. This second hypothesis was tested using a one-sided *t*-test. In summary, proteins higher abundant in ovarian serum than in peripheral serum and additionally higher abundant in ovarian serum from malignant tumors than in ovarian serum from benign tumors were regarded as primary candidates for markers of early stage ovarian cancer. Out of a total of 2,692 spots tested, four spots satisfied the outlined criteria (spots 736, 1,252, 1,295, and 1,507, see **Figures 2A,B**) and were subjected to identification by LC-MS/MS. JUP was identified in all analyzed spots with at least 7 unique peptides (see **Table S3**) and chosen for further analysis by ELISA.

JUP Plasma Concentrations Are Increased in Early Stage Serous Ovarian Cancer

The performance of JUP to distinguish patients with early and advanced stage ovarian cancer from normal controls was assessed using an international multicenter patient cohort.

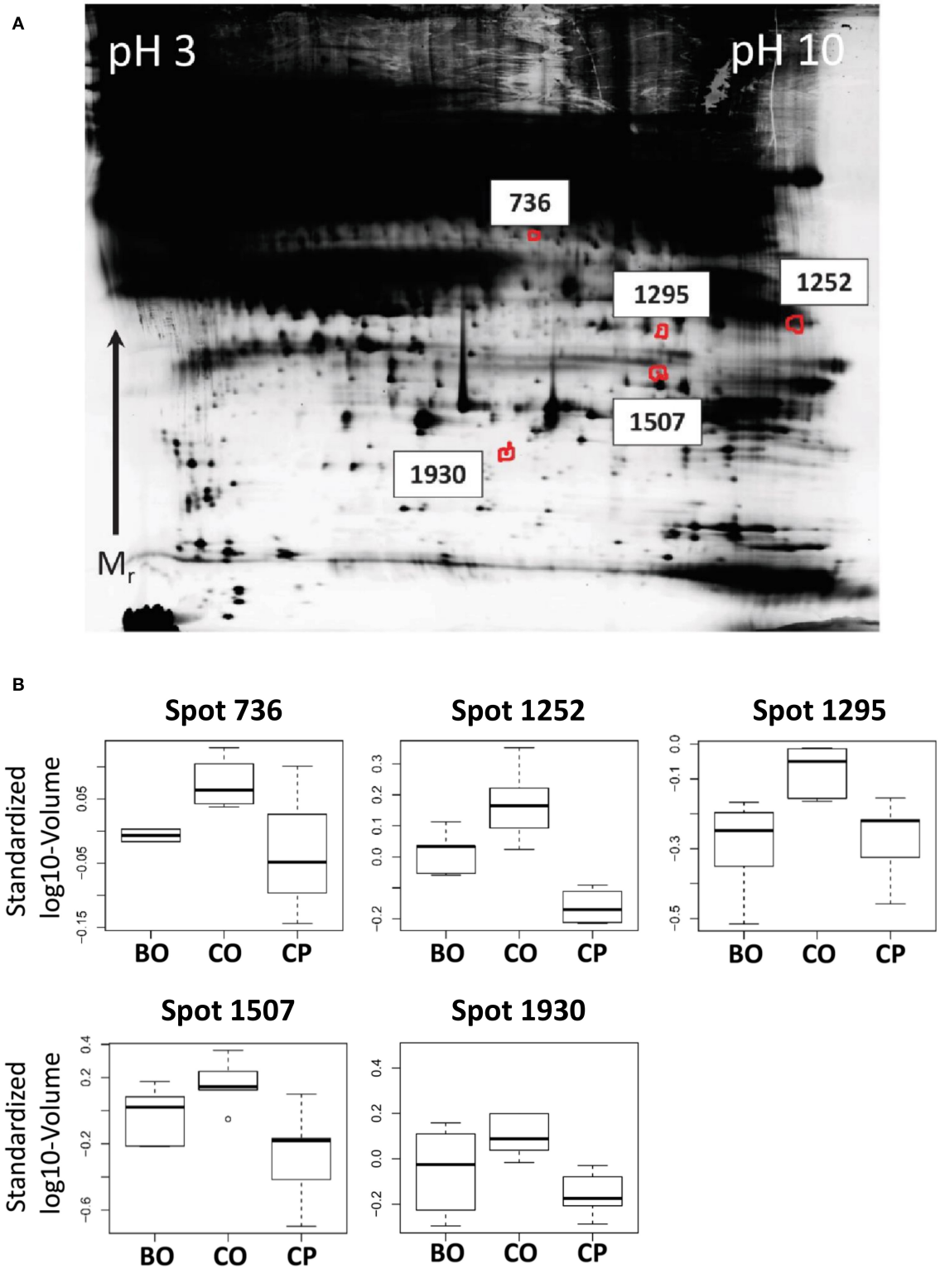


FIGURE 2 | (A) Representative, contrast-adjusted DIGE image of depleted serum. Numbers indicate spots higher abundant in ovarian vs. peripheral blood from early stage ovarian cancer patients and ovarian blood from patients with benign ovarian lesions. **(B)** Boxplots of spot abundance across the analyzed DIGE gels. Samples: CO, Cancer ovarian serum; CP, Cancer peripheral serum; and BO, Benign ovarian serum. All p-values between CO vs. CP and CO vs. BO < 0.05. JUP was identified in all spots by mass spectrometry (see **Table S3**).

Plasma JUP levels were significantly elevated in serous stage IA+B (1.97-fold increase; $p < 0.001$; $n = 20$), serous stage I (2.09-fold increase; $p < 0.0001$; $n = 40$) and advanced serous stage II (1.81-fold increase, $p < 0.001$, $n = 23$), serous stage III ovarian carcinomas (1.98-fold increase; $p < 0.0001$; $n = 34$) and all stages (1.99-fold-change; $p < 0.0001$, $n = 97$) vs. normal controls ($n = 109$, **Figure 3A**). To exclude that observed JUP differences between ovarian cancer patients and controls were caused by an origin from different centers, single center subsets of the data were analyzed. JUP plasma concentrations in samples from the Robinson Research Institute were found to be significantly elevated in stage I ovarian cancer (2.45-fold increase; $p = 0.011$; $n = 8$) vs. normal controls ($n = 82$, **Figure S4**). This was also found in samples from the Sahlgrenska Cancer Center, JUP plasma concentration was elevated in stage IA+B (2.03-fold increase; $p = 0.016$; $n = 9$) and stage I ovarian cancer (1.75-fold increase; $p = 0.008$; $n = 15$) vs. normal controls ($n = 19$; **Figure S4**). These finding make it very unlikely that our results were confounded by center or selection bias. As next step, the performance of JUP was contrasted against CA125 using ROC. JUP exhibited a similar area under the curve (AUC) as CA125 in stage I (JUP: 0.867; CA125: 0.947) and stage IA+B samples (JUP: 0.880; CA125: 0.906) (**Figures 3B,C**). ROC for stage II, III, and combined stage I-III is shown in **Figure S5**. JUP and CA125 were only weakly correlated in ovarian cancer stage I-III ($\rho = 0.327$) as well as in stage IA+B ($\rho = 0.392$) and moderately correlated in stage I ($\rho = 0.410$, data not shown). We therefore established a combined model of JUP and CA125 using logistic regression which resulted in a greater AUC for stage I (0.965) and stage IA+B ovarian cancer (0.941) than for CA125 alone (**Figures 3B,C**). Sensitivity and specificity levels for all markers and models are displayed in **Table 2**.

We assessed if JUP plasma concentrations might be age-dependent, as the average age in the control group was slightly lower. However, the correlation analysis of JUP plasma concentrations and age revealed only a weak correlation in normal controls ($\rho = 0.188$), no correlation in stage I patients ($\rho = -0.014$) and a weak negative correlation in stage IA+B cases ($\rho = -0.382$) (Spearman's rank test) (**Figure S6**).

JUP Serum Concentration Are Increased in Non-serous Epithelial Ovarian Cancer Subtypes

JUP serum concentration were significantly elevated in stage I non-serous ovarian cancer subtypes such as endometrioid (1.93-fold change; $p < 0.001$; $n = 15$), mucinous (1.82-fold change; $p = 0.023$; $n = 12$) and clear cell carcinomas (1.79-fold change; $p = 0.033$; $n = 11$) versus normal controls ($n = 39$, **Figure 4**). ROC analysis was conducted to assess the ability of JUP to distinguish between these early stage cancer subtypes and normal controls. In endometrioid carcinomas JUP exhibited an AUC of 0.886, in mucinous carcinomas 0.748 and in clear cell carcinomas 0.690, which is lower than the AUC values of CA125 (endometrioid: 0.925; mucinous: 0.922; clear cell carcinomas: 0.864). As with the serous ovarian cancer samples, we also conducted a logistic regression analysis to determine the effect of the combining JUP and CA125. This resulted in AUC values similar to those with single markers for all tested non-serous subtypes (endometrioid: 0.905; mucinous: 0.923 and clear cell: 0.834) (**Figure S7**), although there was an increase in AUC for stage IA clear cell carcinoma: JUP: 0.783, CA125: 0.865, JUP+CA125: 0.895 (**Figure S7**). Sensitivity and Specificity levels for all markers, models and subtypes are shown in **Table S8**.

JUP Plasma Concentrations Are Not Increased in Early Stage Breast Cancer

JUP plasma concentrations were not significantly elevated in stage I and II breast cancer (median = 50.4 ng/mL; $p = 0.122$; $n = 12$) when compared to normal controls (median = 36.5 ng/mL; $n = 7$), while JUP plasma levels were significantly increased in stage I and II serous ovarian cancer (median = 68.8 ng/mL; 1.76-fold change; $p = 0.043$; $n = 5$, **Figure S9**). Further analysis of the breast cancer data was conducted nonetheless as an isolated p-value is not a reasonable benchmark for a potential biomarker. ROC analysis was performed and JUP exhibited an AUC of 0.679 in early stage breast cancer (**Figure S9**).

DISCUSSION

Ovarian cancer accounts for about 6% of cancer deaths in women and has the highest mortality of all gynecologic malignancies. The best strategy to reduce mortality would be early detection through screening. However, a reliable early detection test which is able to identify the disease at an early, curable state does not exist.

The to-date largest screening trial, the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), employing an algorithm incorporating the CA125 profile over time, failed to reduce ovarian cancer mortality (30). The US Preventative Services Task Force therefore determined that harms of ovarian cancer screening currently outweigh the benefits and advised against ovarian cancer screening in the general population (31).

Numerous efforts have been undertaken to develop a biomarker-based early detection test [reviewed in (32)]. Proteins that are over-expressed by cancer cells and then released into the bloodstream remain the ideal markers for early detection. However, a major "quantitative" challenge of ovarian cancer biomarker research is to identify the cancer at an early stage when it is small, and the amount of cancer-specific proteins secreted into the blood stream is extremely low (5). To overcome this difficulty, we hypothesized that protein markers are higher concentrated in the ovarian venous blood directly downstream of an ovarian cancer than in peripheral blood taken of the cubital vein where the marker is diluted in the total blood volume. The technique of blood sampling from ovarian veins has been reported for endocrinology studies in the past (33, 34) (**Figure 1**) but to the best of our knowledge, it has never been applied for biomarker discovery in ovarian cancer.

Our strategy resulted in the identification of a promising new marker for epithelial ovarian cancer, JUP. It is primarily considered a protein whose function is to maintain appropriate cell-cell adhesion, but recent studies have also demonstrated

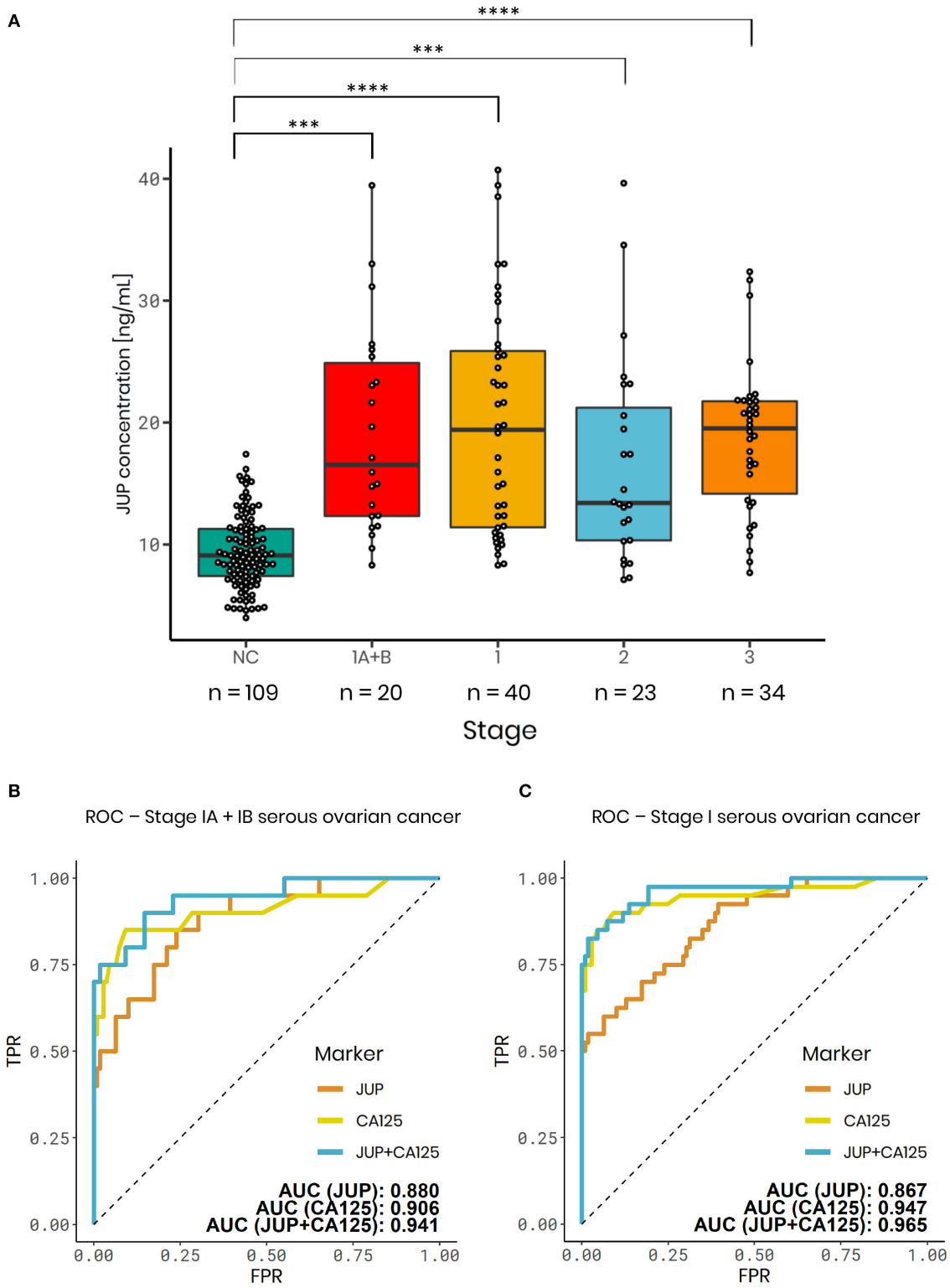


FIGURE 3 | JUP plasma ELISA results for the international, multicenter serous ovarian cancer validation cohort. **(A)** Boxplots of normal controls (NC) (median = 9.09 ng/mL), serous ovarian carcinomas - Stage IA+B (median = 15.45 ng/mL), stage I (median = 18.13 ng/mL), stage II (median = 13.49 ng/mL), and stage III (median = 19.52 ng/mL). Serous carcinomas vs. NC ($n = 109$) - Stage IA+B (1.97-fold change; $p < 0.001$; $n = 20$); stage I (2.09-fold change; $p < 0.0001$; $n = 40$); stage II (1.81-fold change; $p < 0.001$; $n = 23$) and stage III (1.98-fold change; $p < 0.0001$; $n = 34$). **(B,C)** ROC analysis of JUP (orange), CA125 (yellow) and a logistic regression model combining JUP + CA125 (blue). AUC values: **(B)** Serous carcinomas - Stage IA+B: JUP: 0.880; CA125: 0.906; JUP + CA125: 0.941. **(C)** Serous carcinomas - Stage I: JUP: 0.867; CA125: 0.947; JUP + CA125: 0.965. ***p-value < 0.001, ****p-value < 0.0001.

TABLE 2 | Sensitivities at a fixed specificity of 100 and 99.6%, specificities at fixed sensitivity of 75% and area under the curve (AUC) of JUP, CA125 and the combination of both markers for stage I–III serous ovarian cancer.

Model	Stage	Sensitivity (100% specificity)	Sensitivity (99.6% specificity)	Specificity (75% sensitivity)	AUC	PPV
JUP	IA	0.444	0.444	0.807	0.868	0.002
	IA + IB	0.400	0.422	0.826	0.880	0.002
	IC	0.600	0.600	0.688	0.855	0.001
	I	0.500	0.511	0.761	0.867	0.001
	II	0.391	0.410	0.755	0.818	0.001
	III	0.618	0.656	0.945	0.916	0.005
	I–III	0.515	0.538	0.791	0.873	0.001
CA125	IA	0.500	0.524	0.940	0.896	0.005
	IA + IB	0.550	0.572	0.954	0.906	0.007
	IC	0.800	0.844	1.000	0.989	1.000
	I	0.675	0.708	0.991	0.947	0.032
	II	0.783	0.802	1.000	0.971	1.000
	III	0.971	0.983	1.000	1.000	1.000
	I–III	0.804	0.827	1.000	0.971	1.000
JUP+CA125	IA	0.667	0.667	0.945	0.934	0.005
	IA + IB	0.700	0.700	0.982	0.941	0.016
	IC	0.800	0.844	1.000	0.989	1.000
	I	0.750	0.761	1.000	0.965	1.000
	II	0.783	0.802	1.000	0.970	1.000
	III	1.000	1.000	1.000	1.000	1.000
	I–III	0.804	0.822	1.000	0.976	1.000

Positive predictive values (PPV) at fixed sensitivity of 75%, based on an ovarian cancer prevalence of 1:2,500 in the population.

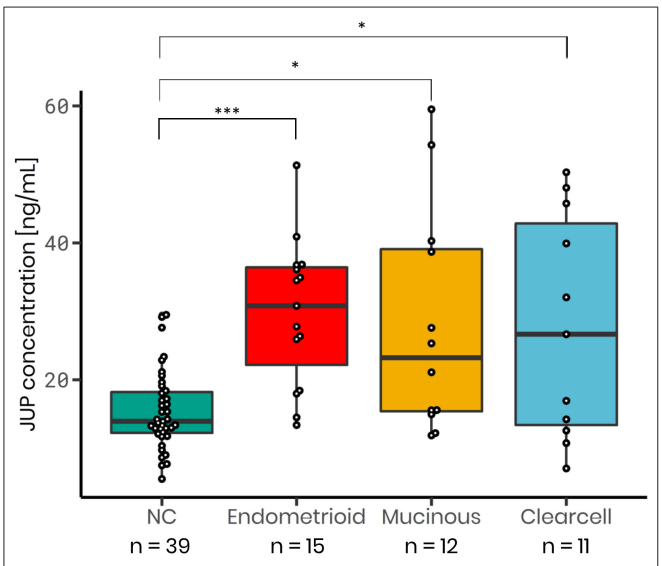


FIGURE 4 | JUP serum concentration is increased in non-serous stage I ovarian cancer subtypes: Endometrioid (median = 30.80 ng/mL; 1.93-fold increase; $p < 0.001$; $n = 15$), mucinous (median = 23.21 ng/mL; 1.82-fold increase; $p = 0.0229$; $n = 12$) and clear cell carcinoma (median = 26.65 ng/mL; 1.79-fold increase; $p = 0.0329$; $n = 11$) as in comparison to normal control samples (NC; median = 13.94 ng/mL; $n = 39$). *P*-values were calculated using a *t*-test. **p*-value < 0.05, ****p*-value < 0.001.

that it is involved in signaling and regulation of tumorigenesis as well as cancer progression (35). In ovarian cancer JUP has mainly been described as tumor suppressor inhibiting *in vitro* growth, migration and invasion (36). It has been shown that JUP interacts with both wild type and mutant p53 and its tumor/metastasis suppressor function in some ovarian cancers may, at least partially, be mediated by this interaction (37). However, JUP's role and regulation on the cellular level remains complex, and both positive and negative roles have been found for various malignancies (38).

JUP was proposed as prognostic biomarker in adenocarcinomas of the lung (39) and as tissue biomarker for cervical (neck) lymph node metastasis in oral squamous cell carcinomas (40). Furthermore, JUP was implicated in testicular germ cell tumors (41), and colorectal cancers (42).

Interestingly, a recent study found that high JUP expression enables tumor cells to adhere together and move as clusters into the bloodstream, facilitating metastasis and resulting in worse prognosis in breast cancer patients (43). It is conceivable that a similar process exists for ovarian cancer, where upregulation of JUP contributes to the formation of ovarian cancer cell aggregates known as spheroids, promoting the release of cancer cells into the abdominal cavity.

Due to the low prevalence of ovarian cancer (1 in 2,500 or 40 in 100,000 postmenopausal women) and the strict requirements for a screening strategy, an effective ovarian cancer screening test

requires a minimum positive predictive value (PPV) of 10%. To achieve a PPV of 10% with a prevalence of 1 in 2,500, a screening test requires a sensitivity of at least 75% for early stage disease and a specificity of at least 99.6%. As the diagnosis of ovarian cancer generally requires a surgical procedure, a PPV of 10% results in 10 operations for every single case of detected cancer (4). In our study, the combination of JUP and CA125 reached a sensitivity of 75 at 100% specificity for FIGO stage I disease and therefore would fulfill the requirements for an early diagnostic test. JUP plasma concentrations were not elevated in early stage breast cancer patients when compared to normal controls, indicating that JUP might be a specific marker for early stage ovarian cancer.

The strength of our research was the inclusion of a large number of stage I serous ovarian cancer cases and the international multicenter validation approach. Limitations are the retrospective nature of the study with analysis of clinical samples in repositories. A larger prospective study is now required to further validate the utility of the biomarker combination for population screening.

In conclusion, our strategy of analyzing ovarian tumor blood for biomarker discovery identified a novel biomarker, JUP, which in combination with CA125 represents a promising novel diagnostic test for early detection of ovarian cancer.

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: PRIDE via ProteomeXchange (http://www.ebi.ac.uk/pride/archive/projects/PXD018417). FTP download: ftp://ftp.pride.ebi.ac.uk/pride/data/archive/2020/08/PXD018417.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee at the Royal Adelaide Hospital, Adelaide, South Australia (Protocol #080304). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FW contributed to the experimental design, performed the proteomics analysis, ELISAs as well as the data analysis, and wrote the manuscript. NL contributed to the experimental design, performed some of the ELISAs, analyzed data, and revised the manuscript. MK-H performed data analysis, helped writing, and revising the manuscript. TJ, AS, and KS acquired and provided ELISA samples and revised the manuscript. PH and MO conceived the project and wrote the manuscript. MO performed the surgery for ovarian blood sampling. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Ovarian Cancer Research Foundation (OCRF), Australia.

ACKNOWLEDGMENTS

The authors want to thank the Ovarian Cancer Research Foundation (OCRF), Australia, for funding of this research, Drs Michelle Turvey, Yin Ying Ho and Georgia Arentz for their help with the ELISAs and Dr. Lyron Winderbaum for his help with the statistical analysis. We are also grateful for the support by the Swedish Cancer Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.01767/full#supplementary-material

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* (2015) 136:E359–86. doi: 10.1002/ijc.29210
2. Maringe C, Walters S, Butler J, Coleman MP, Hacker N, Hanna L, et al. Stage at diagnosis and ovarian cancer survival: evidence from the international cancer benchmarking partnership. *Gynecol Oncol.* (2012) 127:75–82. doi: 10.1016/j.ygyno.2012.06.033
3. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, et al. *Cancer Statistics Review, 1975-2013 - Previous Version - SEER Cancer Statistics Review. SEER.* (2016). Available online at: https://seer.cancer.gov/archive/csr/1975__2013/index.html (accessed January 16, 2020).
4. Clarke-Pearson DL. Clinical practice. Screening for ovarian cancer. *N Engl J Med.* (2009) 361:170–7. doi: 10.1056/NEJMcp0901926
5. Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M. Plasma proteome profiling to assess human health and disease. *Cell Syst.* (2016) 2:185–95. doi: 10.1016/j.cels.2016.02.015
6. Hülsken J, Birchmeier W, Behrens J. E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *J Cell Biol.* (1994) 127:2061–9.
7. Jou TS, Stewart DB, Stappert J, Nelson WJ, Marrs JA. Genetic and biochemical dissection of protein linkages in the cadherin-catenin complex. *Proc Natl Acad Sci USA.* (1995) 92:5067–71.
8. Knudsen KA, Soler AP, Johnson KR, Wheelock MJ. Interaction of alpha-actinin with the cadherin/catenin cell-cell adhesion complex via alpha-catenin. *J Cell Biol.* (1995) 130:67–77.
9. Rimm DL, Koslov ER, Kebriaei P, Cianci CD, Morrow JS. Alpha 1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. *Proc Natl Acad Sci USA.* (1995) 92:8813–7.
10. Witcher LL, Collins R, Puttagunta S, Mechanic SE, Munson M, Gumbiner B, et al. Desmosomal cadherin binding domains of plakoglobin. *J Biol Chem.* (1996) 271:10904–9.
11. Arnold GJ, Fröhlich T. 2D DIGE saturation labeling for minute sample amounts. *Methods Mol Biol.* (2012) 854:89–112. doi: 10.1007/978-1-61779-573-2_7

12. Zammit CM, Weiland F, Brugger J, Wade B, Winderbaum LJ, Nies DH, et al. Proteomic responses to gold(III)-toxicity in the bacterium cupriavidus metallidurans CH34. *Metallomics*. (2016) 8:1204–16. doi: 10.1039/c6mt00142d
13. Dowsey AW, English J, Pennington K, Cotter D, Stuehler K, Marcus K, et al. Examination of 2-DE in the human proteome organisation brain proteome project pilot studies with the new RAIN gel matching technique. *Proteomics.* (2006) 6:5030–47. doi: 10.1002/pmic.200600152
14. Dowsey AW, Dunn MJ, Yang G-Z. Automated image alignment for 2D gel electrophoresis in a high-throughput proteomics pipeline. *Bioinformatics.* (2008) 24:950–7. doi: 10.1093/bioinformatics/btn059
15. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (2014). Available online at: http://www.R-project.org/ (accessed August 26, 2020).
16. Wickham H. The split-apply-combine strategy for data analysis. *J Statis Softw.* (2011) 40:1–29. doi: 10.18637/jss.v040.i01
17. Wickham H. Reshaping data with the reshape package. *J Statist Softw.* (2007) 21:1–20. doi: 10.18637/jss.v021.i12
18. Weiland F, Arentz G, Klingler-Hoffmann M, McCarthy P, Lokman NA, Kaur G, et al. Novel IEF peptide fractionation method reveals a detailed profile of N-terminal acetylation in chemotherapy-responsive and -resistant ovarian cancer cells. *J Proteome Res*. (2016) 15:4073–81. doi: 10.1021/pr100406z
19. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol.* (2012) 30:918–20. doi: 10.1038/nbt.2377
20. Eng JK, Jahan TA, Hoopmann MR. Comet: an open-source MS/MS sequence database search tool. *Proteomics.* (2013) 13:22–4. doi: 10.1002/pmic.201200439
21. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* (2019) 47:D442–50. doi: 10.1093/nar/gky1106
22. Ritz C, Baty F, Streibig JC, Gerhard D. Dose-Response analysis using R. *PLoS ONE.* (2015) 10:e0146021. doi: 10.1371/journal.pone.0146021
23. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag (2009).
24. Clarke E, Sherrill-Mix S. *ggbeeswarm: Categorical Scatter (Violin Point) Plots*. (2017). Available online at: https://CRAN.R-project.org/package=ggbeeswarm
25. Chang W. *extrafont: Tools for Using Fonts.* (2014). Available online at: https://CRAN.R-project.org/package=extrafont (accessed August 26, 2020).
26. Ram K, Wickham H. *wesanderson: A Wes Anderson Palette Generator*. (2018). Available online at: https://CRAN.R-project.org/package=wesanderson (accessed August 26, 2020).
27. Wickham H, François R, Henry L, Müller K. *dplyr: A Grammar of Data Manipulation*. (2019). Available online at: https://CRAN.R-project.org/package=dplyr (accessed August 26, 2020).
28. Bengtsson H. *matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors)*. (2020). Available online at: https://CRAN.R-project.org/package=matrixStats (accessed August 26, 2020).
29. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* (2011) 12:77. doi: 10.1186/1471-2105-12-77
30. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK collaborative trial of ovarian cancer screening (UKCTOCS): a randomised controlled trial. *Lancet.* (2016) 387:945–56. doi: 10.1016/S0140-6736(15)01224-6
31. US Preventive Services Task Force, Grossman DC, Curry SJ, Owens DK, Barry MJ, Davidson KW, et al. Screening for ovarian cancer: US preventive services task force recommendation statement. *JAMA.* (2018) 319:588–94. doi: 10.1001/jama.2017.21926
32. Muinao T, Deka Boruah HP, Pal M. Multi-biomarker panel signature as the key to diagnosis of ovarian cancer. *Heliyon.* (2019) 5:e02826. doi: 10.1016/j.heliyon.2019.e02826
33. Fogle RH, Stanczyk FZ, Zhang X, Paulson RJ. Ovarian androgen production in postmenopausal women. *J Clin Endocrinol Metab.* (2007) 92:3040–3. doi: 10.1210/jc.2007-0581
34. Dickerson RD, Putman MJ, Black ME, Pinto KR, Diamond NG, Marynick S, et al. Selective ovarian vein sampling to localize a leydig cell tumor. *Fertil Steril.* (2005) 84:218. doi: 10.1016/j.fertnstert.2004.12.055
35. Aktary Z, Pasdar M. Plakoglobin: role in tumorigenesis and metastasis. *Int J Cell Biol.* (2012) 2012:189521. doi: 10.1155/2012/189521
36. Alaee M, Danesh G, Pasdar M. Plakoglobin reduces the *in vitro* growth, migration and invasion of ovarian cancer cells expressing N-cadherin and mutant p53. *PLoS ONE.* (2016) 11:e0154323. doi: 10.1371/journal.pone.0154323
37. Aktary Z, Kulak S, Mackey J, Jahroudi N, Pasdar M. Plakoglobin interacts with the transcription factor p53 and regulates the expression of 14-3-3σ. *J Cell Sci.* (2013) 126:3031–42. doi: 10.1242/jcs.120642
38. Aktary Z, Alaee M, Pasdar M. Beyond cell-cell adhesion: plakoglobin and the regulation of tumorigenesis and metastasis. *Oncotarget.* (2017) 8:32270–91. doi: 10.18632/oncotarget.15650
39. He X, Zhou T, Yang G, Fang W, Li Z, Zhan J, et al. The expression of plakoglobin is a potential prognostic biomarker for patients with surgically resected lung adenocarcinoma. *Oncotarget.* (2016) 7:15274–87. doi: 10.18632/oncotarget.7729
40. Kurokawa A, Nagata M, Kitamura N, Noman AA, Ohnishi M, Ohyama T, et al. Diagnostic value of integrin α3, β4, and β5 gene expression levels for the clinical outcome of tongue squamous cell carcinoma. *Cancer.* (2008) 112:1272–81. doi: 10.1002/cncr.23295
41. Skotheim RI, Monni O, Mousses S, Fosså SD, Kallioniemi O-P, Lothe RA, et al. New insights into testicular germ cell tumorigenesis from gene expression profiling. *Cancer Res.* (2002) 62:2359–64.
42. Luque-García JL, Martínez-Torrecuadrada JL, Epifano C, Cañamero M, Babel I, Casal JI. Differential protein expression on the cell surface of colorectal cancer cells associated to tumor metastasis. *Proteomics.* (2010) 10:940–52. doi: 10.1002/pmic.200900441
43. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell.* (2014) 158:1110–22. doi: 10.1016/j.cell.2014.07.013

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 © 2020 Weiland, Lokman, Klingler-Hoffmann, Jobling, Stephens, Sundfeldt, Hoffmann and Oehler. 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 Oncology

ORIGINAL RESEARCH
published: 30 September 2020
doi: 10.3389/fonc.2020.557737

Down-Regulation of Collagen Hydroxylation in Colorectal Liver Metastasis

*Nick A. van Huizen[1,2], Peter C. Burgers[2], Joost van Rosmalen[3], Michail Doukas[4], Jan N. M. IJzermans[1] and Theo M. Luider[2]**

[1] Department of Surgery, Erasmus University Medical Center, Rotterdam, Netherlands, [2] Department of Neurology, Erasmus University Medical Center, Rotterdam, Netherlands, [3] Department of Biostatistics, Erasmus University Medical Center, Rotterdam, Netherlands, [4] Department of Pathology, Erasmus University Medical Center, Rotterdam, Netherlands

Collagen is significantly upregulated in colorectal liver metastasis (CRLM) compared to liver tissue. Expression levels of specific collagen types in CRLM resemble those in colorectal cancer (CRC) and colon tissue. We investigated whether the collagen hydroxylation pattern from the primary tumor also migrates with the metastatic tumor. The degree of collagen alpha-1(I) hydroxylation in colon, CRC, liver, and CRLM tissue of the same individuals ($n = 14$) was studied with mass spectrometry. The degree of hydroxylation was investigated in 36 collagen alpha-1(I) peptides, covering 54% of the triple helical region. The degree of hydroxylation in liver tissue was similar to that in colon tissue. The overall degree of hydroxylation was significantly lower (9 ± 14%) in CRC tissue and also significantly lower (12 ± 22%) in CRLM tissue compared to colon. Furthermore, eleven peptides with a specific number of hydroxylations are significantly different between CRLM and liver tissue; these peptides could be candidates for the detection of CRLM. For one of these eleven peptides, a matching naturally occurring peptide in urine has been identified as being significantly different between patients suffering from CRLM and healthy controls. The hydroxylation pattern in CRLM resembles partly the pattern in liver, primary colorectal cancer and colon.

Keywords: post-translational modification, mass spectrometry, colorectal cancer, colorectal liver metastasis, collagen, hydroxyproline

OPEN ACCESS

Edited by:
Harsha Gowda, The University of Queensland, Australia

Reviewed by:
Nicolò Musso, University of Catania, Italy
Youlang Zhou, Affiliated Hospital of Nantong University, China

***Correspondence:**
Theo M. Luider
t.luider@erasmusmc.nl

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *30 April 2020*
Accepted: *27 August 2020*
Published: *30 September 2020*

Citation:
van Huizen NA, Burgers PC, van Rosmalen J, Doukas M, IJzermans JNM and Luider TM (2020) Down-Regulation of Collagen Hydroxylation in Colorectal Liver Metastasis. Front. Oncol. 10:557737. doi: 10.3389/fonc.2020.557737

INTRODUCTION

Collagen is a family of proteins that form triple helical structures. A triple helix is formed by three protein subunits called alpha-chains. The specific part of an alpha-chain involved in triple helix formation consists of a repeating unit $[\text{-Gly-Xaa-Yaa-}]_n$ where Xaa and Yaa can be different amino acids, frequently proline and lysine (1). Proline and lysine can be enzymatically modified prior to triple helix formation. Proline is hydroxylated into hydroxyproline (Hyp); two different Hyp isomers exist in collagen, namely 4-hydroxyproline (4Hyp) and—more rarely—3-hydroxyproline (3Hyp). Normally, 4Hyp at the Yaa-position is formed by 4-prolyl hydroxylase, which requires Gly-Xaa-Pro as substrate (2). 4-hydroxyproline rarely occurs at the Xaa positions (4xHyp), neither the enzyme, nor the substrate required for the formation of 4xHyp is known (3). 3Hyp is formed by 3-prolyl hydroxylase, which requires Gly-Pro-4Hyp as substrate (4). 3Hyp and 4Hyp are involved in triple helix stability and fiber formation, but their specific functions are still not fully understood (5–7). The function of 4xHyp remains unknown (3). Lysine can be modified into 5-hydroxylysine

(5Hyl) by lysyl hydroxylases (8). In addition, 5Hyl can be further glycosylated (9) and is involved in collagen cross-linking (10). After triple helix formation, collagen triple helices are excreted from the cell via the Golgi apparatus into the extracellular matrix. In the extracellular matrix, triple helices form a variety of supramolecular structures (11).

Triple helix folding is reversible; unfolding occurs when the melting temperature (T_m) is exceeded (12). The T_m is influenced by the physical properties of the amino acids at the Xaa and Yaa positions (1) and post-translational modifications (PTMs). In short model peptides, hydroxylation of proline into 4Hyp increased the T_m up to 6°C, while conversion into 3Hyp decreased the T_m (6). The effect of 5Hyl on the T_m remains unknown. Leikina et al. (13) demonstrate that the T_m of collagen type I in lung is just below body temperature, in addition, only folded triple helices are present below 30°C. If the number of PTM's decreases, then this would result in collagen with a lower Tm. Collagen with a lower Tm will unfold/melt before it can be built into fibers and will be degraded by the body.

The 4Hyp formation rate is influenced by the amino acid at the Xaa position (14), the temperature at which collagen-producing cells are cultured (15), and the amount of oxygen present (16). Collagen produced under hypoxic conditions has ∼1.7% less Hyp and almost 10% less Hyl (16). Under hypoxic conditions, proline hydroxylation is the rate limiting step of collagen formation, indicating that a minimum number of PTMs is required to form stable collagen. Because collagen hydroxylation is influenced by many parameters, we hypothesize that the hydroxylation pattern varies strongly between different tissues and within a particular tissue. In agreement with this, Montgomery et al. (17) described that the degree of hydroxylation in collagen alpha-1(I) of ten breast cancer patients varied between 0 and 90%, depending on the proline position.

Collagen levels in tumor tissue are dysregulated in comparison to healthy tissue (18). In colorectal liver metastasis (CRLM), collagen levels are, in general, significantly upregulated in comparison to adjacent normal liver tissue (19, 20). Collagen type 10, 12, 14, and 15 expression levels have most likely migrated from the healthy colon tissue, via the primary tumor (colorectal cancer; CRC), to the metastasis. The migration of colon-specific protein expression into CRLM is not limited to collagen; other proteins, e.g., CDH17, KRT20, CEACAM5, GPA33, MUC13, and PPP1R1B/DARPP32 show a similar pattern, although many colon-specific proteins are down-regulated too (20).

In this study, we have investigated the hydroxylation pattern, taking into account the hydroxylation of proline and lysine in collagen alpha-1(I), to better understand the changes in collagen caused by CRLM. We hypothesize that the hydroxylation pattern co-migrates from colon tissue, via the primary tumor, toward the metastasis similarly as a number of colon-specific proteins and specific collagen types. In addition, it was investigated if studying the collagen hydroxylation pattern can aid in the finding of biomarkers to further improve the diagnoses of CRLM based on urine (21). Collagen hydroxylation was studied with bottom-up mass spectrometry in colon, CRC, CRLM, and liver tissue obtained from the same individual.

Abbreviations: 3Hyp, 3-hydroxyproline; 4Hyp, 4-hydroxyproline; 4xHyp, 4-hydroxyproline at the Xaa-position; CRC, Colorectal cancer; CRLM, Colorectal liver metastasis; Hyp, Hydroxyproline; IQR, Interquartile range; NOP, Natural occurring peptides; pO_2, Partial pressure of oxygen; T_m, Melting temperature of collagen.

MATERIALS AND METHODS

Sample Selection

Bottom-up proteomics data of matching normal liver tissue and colorectal liver metastasis (CRLM) tissue of 30 patients were retrieved from the PRIDE Archive (PXD008383). Normal liver tissue was taken at least 1 cm away from the tumor (20). We could collect matching healthy adjacent colon tissue and CRC tissue from 14 out of these 30 patients. The tissue material had been collected during curative surgery. Morphology differences indicated that the CRC tissues contained adjacent normal tissue, and that one colon tissue contained tumorous tissue. None of the CRLM tissues contained non-tumor material, and none of the liver tissues contained non-normal tissue. This study was approved by the Erasmus MC ethics review board (MEC-2007-088) and we have worked according to the declaration of Helsinki. An experienced pathologist (MD) has reviewed the sections.

Sample Preparation and Data Processing

All tissues were prepared, measured, and data processed as described by van Huizen et al. (20). In short, formalin cross-links were removed by incubation with 1M TRIS pH 8, followed by cysteine reduction and alkylation. The proteins in the tissue were cleaved into peptides by overnight incubation with trypsin, 0.6 μg of trypsin was added. Trypsin cleaves after lysine and arginine except if a proline is present after these two amino acids. Peptides were identified with nanoLC-MS/MS measurements on an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Fischer Scientific, San Jose, CA, USA). MGF peak list files were extracted from raw files by ProteoWizard (v3.0.9166) and searched with Mascot (v2.3.2, Matrix Science Inc., London, UK) and the UniProt/Swiss-prot Human database (20,194 entries, v2015_11). Hydroxylations (+16 Da) of proline, lysine, and methionine were included as variable modifications and carbamidomethylation (+57 Da) of cysteine as fixed modification. A maximum of 4 missed cleavages was allowed. Mascot search results were further analyzed in Scaffold (v4.6.2, Portland, OR, USA). In Scaffold, protein confidence levels were set to a 1% false discovery rate (FDR), at least 2 peptides per protein, and a 1% FDR at the peptide level. FDRs were estimated by inclusion of a decoy database search generated by Mascot. A Peptide Report exported from Scaffold was used for data analysis. The peptide report regarding liver and CRLM tissue has been published (20), the peptide report regarding colon and CRC tissue is included in the **Supplementary S1**. Mass spectrometry data was made publicly accessible via the PRIDE archive, accession number: PXD015015.

Peptide Panel Selection

Collagen alpha-1(I) peptides were selected on the following criteria: (1) does not contain methionine or cysteine; (2) contains

at least one proline or lysine moiety; and (3) a unique number of hydroxylations, having been identified in 20 or more liver or CRLM tissues, and in 3 or more colon and CRC tissues. The latter sub criterion was applied because the statistical formulas used require a sample size of 3 or larger.

Principal Components Analysis

Principal component analysis (PCA) was performed in R (v3.6.0.) (22); the R-script can be found in the **Supplementary S2**. A table was created containing for every sample a value for the absence (value 0) or presence (value 1) for each peptide with a unique number of hydroxylations. Zero to 10 hydroxylations were taken into consideration; where 10 is the maximum number of hydroxylations observed for the peptides in the peptide panel. The PCA was calculated with the "prcomp" function. If two or more tissue types were compared, then a 2D-PCA-plot was made with the package "factoextra." 3D-PCA-plots were made by exporting the coordinates of the first three PCA dimensions, and plotting with "scatter3d." For assessment of the data, 2D-PCA-plots were made by exporting the PCA coordinates, and plotting with "ggplot2."

Hydroxylation Analysis

Statistical analysis was performed in R (v3.6.0.) (22), the R-script can be found in the **Supplementary S3**. In the digested tissue, collagen peptides are present with different numbers of hydroxylations per peptide. A difficulty with collagen analysis is the identification of the exact location of a hydroxylation. To avoid the identification of the exact hydroxylation position, we used the number of hydroxylations per peptide obtained by the parent mass. Per tissue type ($n = 4$), per patient ($n = 14$ and $n = 30$), and per peptide from the peptide panel ($n = 36$), we checked the presence of 0 to 10 hydroxylations taking into account both hydroxyproline and hydroxylysine, which generated a three-dimensional data matrix. Per peptide a list was generated containing the numbers of different forms observed, compared between two tissue types. Testing for differences between tissue types was performed with a paired Wilcoxon signed-rank test. Per tissue a list was generated containing the average number of hydroxylations per peptide, which was compared between two tissue types with a paired Wilcoxon signed-rank test. If a peptide was not identified in tissue from a particular patient, no average could be calculated, then this patient was removed from analysis of the average number of hydroxylations. Besides the overall differences between the number of different forms observed, and the average number of hydroxylations per peptide, we also studied in more detail if there are peptides with a specific number of hydroxylations that is different between different tissues. The difference in a specific hydroxylation form between two tissue types was assessed with the McNemar test. Two-sided p-values below 0.05 were considered as significant. **Figure 1** shows an example of the set-up of the data.

Data Assessment

The quality of the data was assessed prior to the analysis of collagen hydroxylation. A bias in the data based on patient characteristics (age and gender) was investigated by making 2D-PCA plots per tissue type. For cancerous tissue also the location of the primary tumor was taken into account. Furthermore, colon and CRC tissue contained non-normal and non-tumor tissue, respectively. Therefore, an additional 2D-PCA-plot was made containing colon and CRC tissue to assess the presence of a bias based on the percentage of non-normal and non-tumor tissue.

Permutation testing (i.e., randomly reassigning samples to the groups) was used to estimate background in the distribution of the number of forms, the average number of hydroxylations and specific hydroxylation forms between different tissue types. For each peptide, the distribution of the number of forms, the average number of hydroxylations and specific hydroxylation forms was compared between liver and CRLM (both $n = 30$ and $n = 14$) and between other pairs of organs ($n = 14$) using a paired Wilcoxon signed-rank test with a critical p-value of 0.05. To demonstrate the significant differences present between the unpermitted data (i.e., the true data), the permutation results were used as a test statistic. The critical value of the permutation test was calculated as the 95th percentile of the number of peptides with a significant difference when samples were permutated. The significance of the number of differences found in the unpermitted data was expressed as a p-value. This p-value was calculated as the proportion of the permutations where the number of peptides with a significant difference was at least as large as in the unpermitted data. The test used 1,000 random permutations.

Presence of linear relation between the number of forms and the average number of hydroxylations was tested per tissue with the Pearson correlation test. A linear relation was considered significant if the p-value was below $0.05/4 = 0.0125$ (Bonferroni correction for multiple testing, $n = 4$).

Analysis of the Hydroxylation Variation

The hydroxylation pattern was compared between: colon and liver, colon and CRC, CRC, and CRLM, colon and CRLM, and liver and CRLM. The difference in the degree of hydroxylation was analyzed in different directions: (1) 2D- and 3D-PCA-plots; (2) the number of forms observed; (3) detailed analyzes of the average number of hydroxylations; and (4) analyses of a specific hydroxylation form. Detailed information regarding the average number of hydroxylations was obtained by: plotting the average number of hydroxylations per peptide; the absolute difference in the average number of hydroxylations per peptide of CRC, CRLM, and liver tissue with respect to colon tissue (tissue—colon tissue); and the relative difference with respect to colon tissue ([tissue/colon −1]*100%). Differences between all groups were tested with independent samples t-tests. For the average number of hydroxylations per peptide, p-values below 0.05/6 (Bonferroni correction, 6 comparisons), and for the relative differences, p-values below 0.05/3 (Bonferroni correction, 3 comparisons) were considered significant. The average of the relative difference was compared to zero with a one-sample t-test; p-values below 0.05/3 (Bonferroni correction, 3 comparisons) were considered as significant.

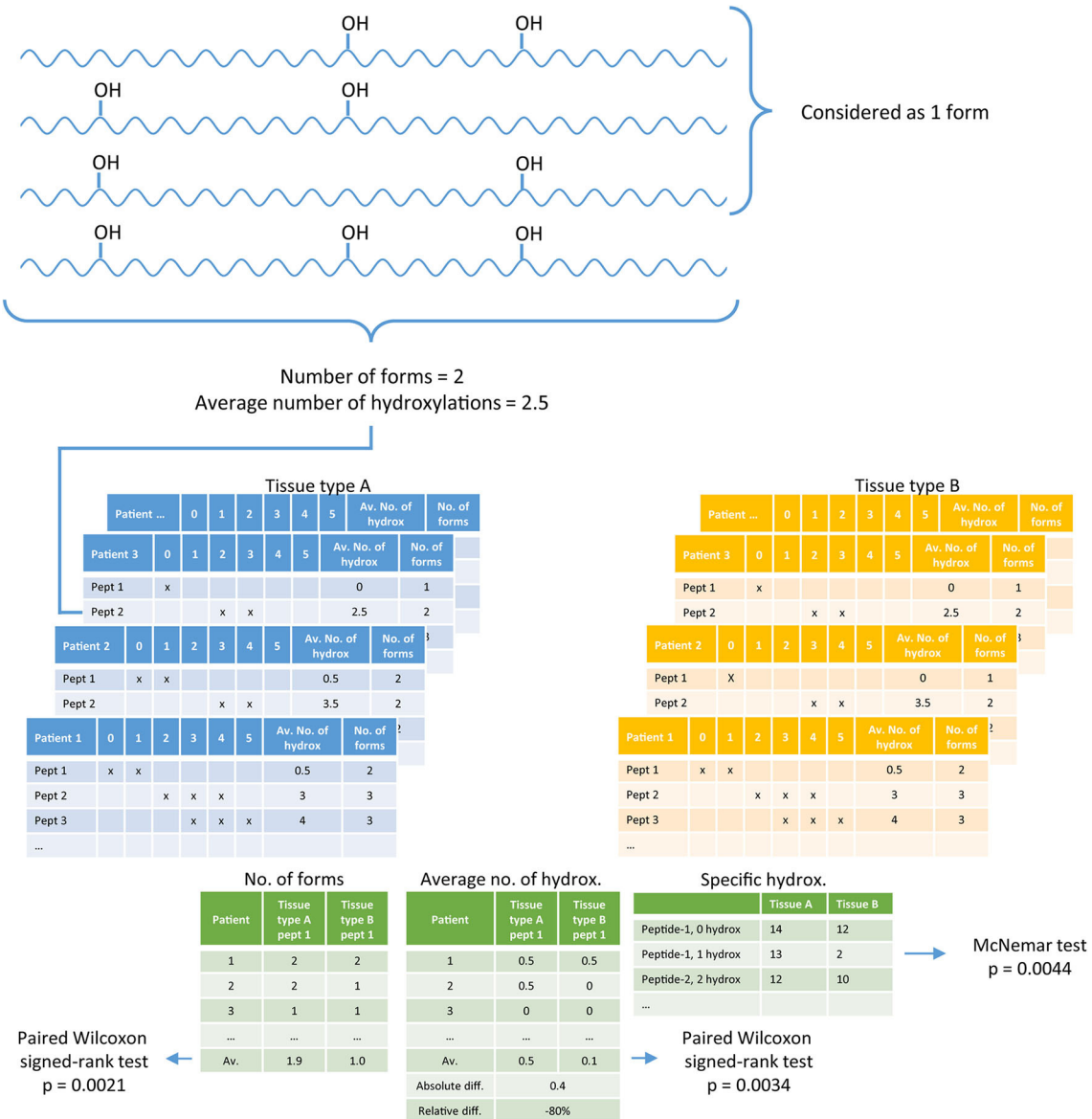


FIGURE 1 | A diagram explaining the data analysis. Per tissue, per patient the number of forms present in collagen peptides, and the average number of hydroxylations per peptide is determined (blue and yellow tables). Per peptide, per patient these two values are compared between two tissue types, left and middle green table. In addition, every form (specific hydroxylation) is separately compared between two tissue types based on the number of patients in which the specific form is identified, right green table.

Urine Analysis

The value of the significant different peptides identified in section Analysis of the Hydroxylation Variation as a biomarker has been investigated by reanalyzing the published data (PXD013533) by van Huizen et al. (21). The bottom-up mass spectrometry data was processed as described by van Huizen et al., whereby we considered a peptide as significant when a *p*-value was below 0.05.

RESULTS

Patient Characteristics

An overview of all patient characteristics is present in **Supplementary S4**. CRC patients ($n = 14$) had a median age of 64 years [interquartile range [IQR] 57–73] and were mostly male (64%). The CRC tumor had a median size of 4.0 cm (IQR 3.3–4.8), being moderately differentiated adenocarcinomas. CRC was located in rectum (64%, 9 out of 14) or sigmoid (36%, 5 out of 14). Liver and CRLM tissue were not collected simultaneously with colon and CRC tissue. The CRLM patients had a median age of 65 years (IQR 59–73) and a median of 2 (maximum 7) tumors with a median size of 2.3 cm (IQR 1.2–3.5), being moderately differentiated adenocarcinomas.

Prior to surgical removal of CRC or CRLM some patients have received pretreatment in the form of chemotherapy (CRC = 0, CRLM = 6), or a combination of chemotherapy and radiotherapy (CRC = 2, CRLM = 1). In **Supplementary S5**, we

have investigated if a clustering based on the pretreatment is present in a 2D-PCA-plot, and we did not observe a correlation.

In some patients molecular diagnostics has been performed on the CRC ($n = 7$) or CRLM ($n = 4$). Only in three of the seven CRC tumor tissues on which molecular diagnostics has been performed are included in this research ($n = 3$). In **Supplementary S5**, we have investigated if a clustering in 2D-PCA-plot based on the results of the molecular diagnostics is present, and we did not observe a correlation.

Peptide Selection

All tissue samples were digested with trypsin and measured via a proteomics bottom-up approach. Mass spectra were searched against the Uniprot/Swissprot human protein database, having hydroxylation of proline and lysine as a variable modification. The database search resulted in the significant identification of, among others, collagen alpha-1(I) peptides.

The number of identified collagen alpha-1(I) peptides with unique numbers of hydroxylation differed per tissue type (ANOVA $p < 0.001$), only colon and CRC were not significantly different (Bonferroni *post-hoc* testing $p > 0.05$) and all other comparisons were significantly different (Bonferroni *post-hoc* testing $p < 0.0001$). The following average numbers of peptides with unique number of hydroxylations were identified in liver 70 ($n = 30$, SD=17), CRLM 111 ($n = 30$, SD = 23), colon 162 ($n = 14$, SD = 10), and CRC 168 ($n = 14$, SD = 28). A total of 113 different peptides with a unique number of hydroxylations were identified, which did neither contain methionine nor cysteine. From these 113 peptides a panel of peptides was selected based on the following selection criteria: (1) does not contain methionine or cysteine; (2) contains at least one proline or lysine moiety; (3) a specific number of hydroxylations in a specific collagen peptide ought to be identified in 20 or more liver or CRLM tissues, and in 3 or more colon and CRC tissues. Thirty-six peptides fulfilled all selection criteria These 36 peptides covered 54% of the triple helical region of collagen alpha-1(I), **Supplementary S5**. They are ranked corresponding to their occurrence in the primary structure of collagen, and will be referred in the text as, for example, peptide 1. **Supplementary S5** also contains per tissue a table specifying the number of hydroxylations on the peptides present.

Data Assessment

Prior to analysis of the collagen hydroxylation pattern, the data was assessed to investigate if any bias is present in the data that would indicate that the degree of hydroxylation is related to a patient characteristic. Data assessment was performed by making 2D-PCA-plots and a permutation test.

In the constructed 2D-PCA-plots we investigated if a relation between the collagen hydroxylation and patient characteristics was present (**Supplementary S5**). The 2D-PCA-plot is based on the absence or presence of a peptide with a specific number of hydroxylations. The patient characteristics are visualized by specific shapes, colors, and different sizes of the data points. The axes of the 2D-PCA-plots describe the first and second dimension variability of the first two principal components (PC).

The degree of collagen hydroxylation in CRLM and CRC tissue showed no clustering based on patient age, gender, or location of the (primary) tumor. Likewise, the degree of hydroxylation in liver and colon tissue showed, based on 2D-PCA-plots, no clustering based on patient age or gender. All CRC tissues contained to some extent normal tissue, and just one colon tissue contained some tumorous tissue. Nevertheless, colon and CRC samples still clustered based on the tissue type. The 2D-PCA-plots are available in the **Supplementary S5**.

Permutation testing results are presented in **Figure 2**. The permutation test indicates that, differences between two tissue types are significant ($p < 0.05$) if more than 4 differences in the number of forms, 3 differences between the average number of hydroxylations per peptide, or 4 specific hydroxylation forms are found. In all comparisons more significant differences were present than expected by permutation testing. The within tissue variation in liver and CRLM tissue is similar to the number of differences observed by chance between the different tissue comparisons, indicating that the differences between tissue are significantly larger than within a tissue.

Tissue Comparison

In the 2D-PCA-plot shown in **Figure 3**, liver, CRLM, and colon are separated based on their hydroxylation pattern. CRC is separated from liver and CRLM tissue, but overlaps with colon tissue. Two exceptions are observed. Patient CRLM-10 is located (coordinates: −3.98, 0.22) in between the liver samples, most likely because only 23 different peptides with a unique number of hydroxylations were identified (liver average = 62). Patient CRC-19 is located (coordinates: −1.01, 0.38) in between the liver samples, because only 69 different peptides with a unique number of hydroxylations were identified (CRC average = 92). A 3D-PCA-plot based on the first three dimensions results in a complete separation between liver, CRLM, and Colon+CRC, while colon and CRC are still partly overlapping. The 3D-PCA-plot is available in the **Supplementary S5**. The 2D-PCA-plots and 3D-PCA-plots for individual tissue comparisons are also available in the **Supplementary S5**. For the comparison "liver vs. colon," there is a complete separation in a 2D-PCA-plot and 3D-PCA-plot. For "colon vs. CRC" there is a partial overlap in the 2D-PCA-plot and 3D-PCA-plot. For "CRC vs. CRLM," and "liver vs. CRLM" (14 vs. 14) there is a slight overlap in the 2D-PCA-plot and they are fully separated in the 3D-PCA-plot. "Liver vs. CRLM" (30 vs. 30) are partly overlapping in the 2D-PCA-plot, and fully separated in the 3D-PCA-plot.

An overview of the results of the tissue comparisons based on the number of forms, average number of hydroxylations, and specific hydroxylations is presented in **Figure 2**. A Pearson correlation test was performed to exclude a relation between the number of different forms observed and the average number of hydroxylations per peptide. Significant correlations were not found (p-critical = 0.05/4 = 0.0125, Bonferroni correction); liver: $r = 0.40$, $p = 0.018$ CRLM: $r = 0.31$, $p = 0.066$; colon: $r = 0.21$, $p = 0.21$; CRC: $r = 0.16$, $p = 0.35$.

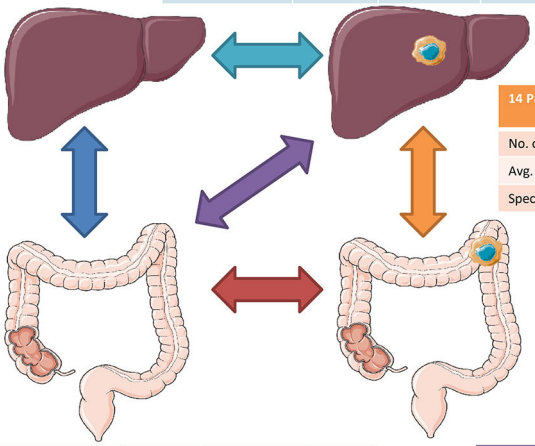
	No. of forms Down	up	Avg. No. of hydrox Down	up	Specific hydrox. Down	Up
Liver vs colon	2	26	8	5	3	27
Colon vs CRLM	18	3	12	1	22	5
Colon vs CRC	4	3	11	0	7	3
CRC vs CRLM	3	8	4	7	13	5
Liver vs CRLM (30)	2	22	19	3	7	30
Liver vs CRLM (14)	1	5	14	0	3	8

30 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	24	5	≈0
Avg. No. of hydrox	22	4	≈0
Specific hydrox	37	6	≈0

14 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	6	3	0.005
Avg. No. of hydrox	14	2	≈0
Specific hydrox	11	3	≈0

14 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	28	3	≈0
Avg. No. of hydrox	13	3	≈0
Specific hydrox	30	3	≈0

14 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	11	3	≈0
Avg. No. of hydrox	11	3	≈0
Specific hydrox	17	3	≈0



14 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	7	3	0.001
Avg. No. of hydrox	11	2	≈0
Specific hydrox	18	2	≈0

14 Patients	Significant (p<0.05)	Permutation based cut-off	P-value
No. of forms	21	4	≈0
Avg. No. of hydrox	13	3	≈0
Specific hydrox	27	3	≈0

FIGURE 2 | Tissue comparison analysis. The black table shows the number of peptides down- or upregulation regarding the differences in the number of forms, the average number of hydroxylations, and specific hydroxylations. The arrows indicate the comparisons made and tables surrounding the tissue cartoons having the same color belong together. Reprinted (adapted) with permission from (21). Copyright (2019) American Chemical Society.

The number of significant peptides and the direction (down- or upregulated) indicate the similarity between the different tissue types. It is worth noting that colon and lover tissue show the highest difference in the average number of hydroxylations, although it is equally down- and upregulated. While colon and CRC have only 11 significant differences in the average number of hydroxylations, the degree of hydroxylation is strongly downregulated in CRC tissue. Also between CRC and CRLM only 11 significant differences were present in the average number of hydroxylations, however in this case they are more equally down- and upregulated.

To compare the effect of group size, the comparison "liver vs. CRLM" was performed with the patients of whom colon and CRC tissue was collected (14 vs. 14) and with the full data set (30 vs. 30). As expected, in the full data set, more and stronger significant differences were observed, because more samples result in a better powered analysis.

In all tissue comparisons the total number of significant differences with respect to the average number of hydroxylations is similar. However, the comparisons of "liver vs. colon," and "CRC vs. CRLM" have a similar number of down- and upregulated averages (black table **Figure 3**), while the comparisons "colon vs. CRC," "colon vs. CRLM," and "liver vs. CRLM" have an overall increase in downregulation of the average.

Figure 4 shows a more detailed analysis of the average number of hydroxylations. In **Figure 4A**, the average number of hydroxylations per peptide per tissue is shown. The overall average numbers of hydroxylations in CRC and CRLM are significantly lower than those in colon and liver tissue. Liver and colon tissue do not differ in number of hydroxylations, and also CRC and CRLM do not differ in this respect. **Figure 4B** shows the differences in the average number of hydroxylations in the comparison of colon tissue and the other tissues investigated. **Figure 4B** visualizes the overall shift in the average number of hydroxylations per peptide. The overall shift in the number of hydroxylations per peptide in liver tissue, with respect to colon tissue, does not significantly differ from zero. Overall, the mean of all average number of hydroxylations per peptide is significantly downregulated in CRC ($p = 6.3^{*}10^{-3}$) and in CRLM ($p = 6.3^{*}10^{-5}$) tissue compared to liver tissue. CRC and CRLM do not have a different shift in comparison to each other ($p = 0.27$). In **Figure 4C**, the relative differences in the average with respect to colon tissue are shown. There is a significant decrease in the average degree of hydroxylation, on average −9% in CRC (SD = 14%, p

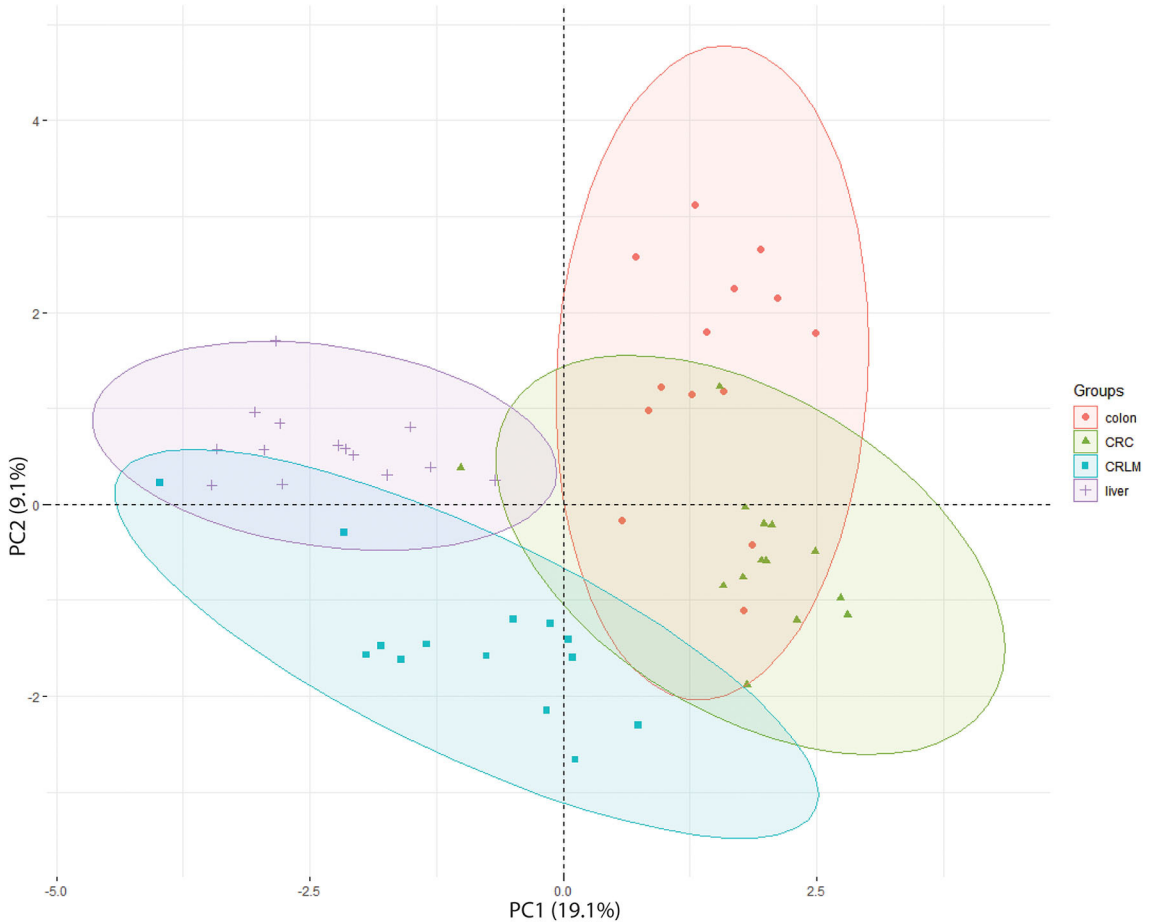


FIGURE 3 | 2D-PCA-plot showing the separation of all 4 tissue types based on their degree of hydroxylation. The x-axis is the first principal component (PC1) of the PCA, which explains 19.1% of the variation. The y-axis is the second principal component (PC2) of the PCA, which explains 9.1% of the variation.

= 0.00056), and on average −12% in CRLM (SD = 22%, p = 0.0026) with respect to colon tissue, and liver tissue does not significantly differ from colon tissue (average = −3%, SD = 27%, p = 0.55).

Figure 4C visualizes how the degree of hydroxylation changes per peptide from colon to CRC, to CRLM, to liver tissue. For peptide 1, the average number of hydroxylations remains constant with the transition from colon to CRC tissue. However, in CRLM tissue the average number of hydroxylations decreases strongly. The degree of hydroxylation of peptide 1 in CRLM is similar to that in liver tissue. Regarding peptide 21, the average number of hydroxylations decreases strongly from colon to CRC, and from CRC to CRLM, and is fairly similar between CRLM and liver tissue. Regarding peptides 15 and 16, the average number of hydroxylations decreases from colon to CRC, and increases from CRC to CRLM to liver tissue. Peptides 1, 15, 16, and 21 appear to move toward liver tissue, when transitioning from colon, via CRC, to CRLM.

The largest number (n = 30) of different specific hydroxylation sites were detected between liver and colon tissue; a comparable number (n = 27) of different specific hydroxylation sites is present between colon and CRLM tissue. A closer study of specific hydroxylations that differ between liver and CRLM tissue, shows several unique differences, see **Table 1**. Peptide 7 with 0 hydroxylations is more prevalent in CRC and CRLM than in healthy colon and liver tissue. Peptide 13 with 1, peptide 15 with 1, peptide 16 with 1, peptide 18 with 0, peptide 19 with 3, and peptide 34 with 2 hydroxylations are abundantly present in colon, CRC and CRLM, but hardly in liver tissue. Peptide sequences are presented in the **Supplementary S5, S6**. Peptide 17 with 2, and peptide 27 with 5 hydroxylation are less abundant in CRLM compared to colon, CRC, and liver. Peptide 14 with 6, and peptide 24 with 5 hydroxylations, are present at similar levels in colon and liver, less in CRC, and the least in CRLM. The general pattern emerging from this data is consistent with a lower degree of hydroxylation in CRLM than in liver or colon tissue.

Urine Analysis

In urine, natural occurring peptides (NOPs) were identified that overlap with six peptides from 1 (21). Only for peptide 34,

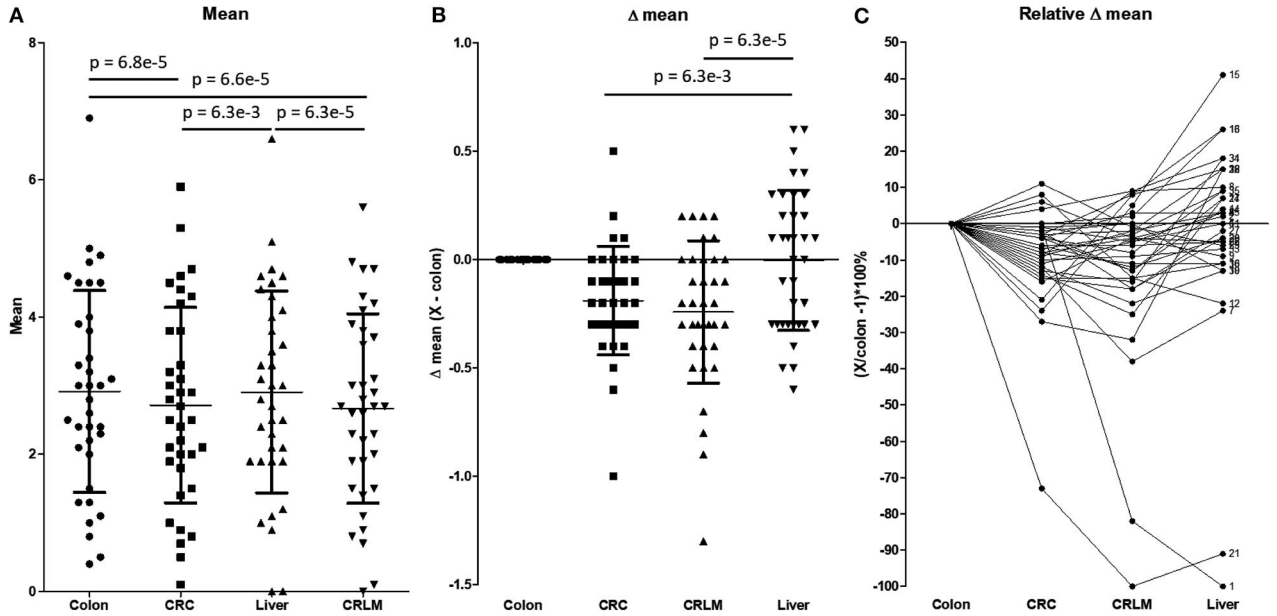


FIGURE 4 | (A) The average number of hydroxylations per peptide. Tested for differences with a repeated paired *t*-test. *P*-values below 0.05/6 = 0.0083 are considered significant. **(B)** Absolute difference in the average number of hydroxylations per peptide with respect to colon tissue. *P*-values below 0.05/3 = 0.0167 are considered significant. **(C)** Relative difference in the number of hydroxylations with respect to colon tissue. The "X" in the legend of the Y-axis represents CRC, CRLM, or liver tissue, depending on the comparison made. Only significant values are presented as horizontal bars in **(A)** and **(B)**.

TABLE 1 | Peptides with a specific hydroxylation form that are significantly different between liver and CRLM tissue, and the number of patients, per tissue, in whom the peptide was identified.

Peptide	Number of prolines and lysines	Number of hydroxylations	14 Patients (matching colon, CRC, liver, and CRLM)				30 Patients (matching liver, and CRLM)	
			Colon	CRC	CRLM	Liver	CRLM	Liver
7	7	0	0	2	8	0	15	0
13	4	1	14	14	12	3	26	13
14	8	6	13	7	3	14	5	27
15	5	1	14	14	12	3	26	6
16	9	2	13	13	8	1	20	3
17	4	2	14	10	4	11	10	18
18	4	0	7	9	11	1	26	3
19	6	3	8	11	11	0	22	3
24	9	5	5	3	0	8	2	13
27	9	5	13	12	3	9	5	17
34	9	2	12	11	10	0	20	2

from 1, a NOP (AGPPGAPGAPGAPGPVGPAGKSGDRGETGP with 2 hydroxylations) was significantly different (*p*-value of $3.1*10^{-4}$) between urine from patients suffering from CRLM and healthy control urine. On the basis of the b- and y-ions from MS/MS spectra, the peptide AGPP(-OH)GAP(-OH)GAPGAPGPVGPAGKSGDRGETGP is identified to be hydroxylated at the 5th and 8th amino acid in colon tissue, CRC tissue, liver tissue, CRLM tissue, control urine, and CRLM urine.

DISCUSSION

The hydroxylation pattern of collagen alpha-1(I) is tissue-specific and each studied tissue has a unique hydroxylation pattern and a degree of hydroxylation as could be detected by PCA and statistical analysis. The variation in the hydroxylation pattern between liver and colon is larger than the variation within each tissue. Even though the average number of hydroxylations is comparable between colon and liver tissue, locally along

the collagen alpha-1(I) chain the degree of hydroxylation is significantly different. The variation within a tissue and between healthy tissues implicates that every collagen triple helix has a unique hydroxylation pattern. Further studies are required to investigate whether prolines at specific sites are always hydroxylated. In addition, it would be of interest to ascertain whether a lack of specific hydroxyprolines could inhibit triple helix formation, or prevent incorporation of such a triple helix into supra-molecular structures. The hydroxylation pattern has a strong influence on the T_m. We suggest that every triple helix consequently has a slightly different T_m, and therefore interacts differently with the surroundings, and has different micro-unfoldings. In previous research it has been assumed that all prolines are hydroxylated; for example, to determine the T_m *in silico* (23). We show that this is not the case and that this assumption leads to an overestimation of the T_m.

The collagen hydroxylation pattern is studied based on 36 selected peptides. Several of these selected peptides contain a missed cleavage present at arginine at lysine. A possible explanation is the presence of hydroxylation at the lysine moiety, which has been described to reduce the activity of trypsin (24). Although, peptide 1, which only contains one lysine moiety, is observed with and without a hydroxylation at lysine. This example shows that trypsin is able to cleave at a hydroxylated lysine. It might be more likely that the presence of aspartic acid and glutamic acid in close proximity to an arginine and lysine moiety reduces the efficiency of trypsin activity (25, 26).

We investigated if the hydroxylation pattern in a tumor migrates with the metastasis. The results of this study indicate that the hydroxylation degree in tumor tissue decreases significantly in comparison to degree of hydroxylation in healthy tissue. The degree of hydroxylation in CRC tissue is 9% lower and that in CRLM is 12% lower than in colon and liver tissue. Still, in the comparisons "colon vs. liver" and "CRLM vs. CRC" similar numbers of peptides were down- and upregulated. The hydroxylation pattern of several of the studied peptides appears to be tissue-dependent. For instance, peptide 1 was present in colon tissue with 0 and 1 hydroxylation, but in liver tissue only with 0 hydroxylation. CRLM tissue seems to be in an intermediate state, containing both peptide 1 with 0 hydroxylation and—to a lesser extent—peptide 1 with 1 hydroxylation. Peptide 1 contains just one lysine moiety, and zero proline moieties, further modification of 5-hydroxylysine (e.g., glycosylation) can prevent detection of peptide 1 with 1 hydroxylation. To our knowledge additional modifications of this specific lysine are not described in literature; nevertheless, if an additional modification is present, then this still indicates a significant difference between the tissues that have been compared. In line with the peptide 1 pattern, the hydroxylation degree of peptide 21 in colon tissue is similar to that in CRC tissue, and strongly reduced in both CRLM and liver tissue. The hydroxylation degree of peptides 15 and 16 decreases in CRC tissue, increases in CRLM tissue, and is even higher in liver tissue. Despite the reduced degree of hydroxylation in CRLM tissue, the general hydroxylation pattern does partly resemble liver tissue but also CRC, and colon.

A recently described method to detect CRLM is by measuring collagen natural occurring peptides (NOPs) in urine (27, 28). Higher sensitivity and specificity can potentially be obtained by adding distinctive collagen NOPs to the panel of molecular markers as described by us. In the present study, we found an overall reduction in the degree of collagen hydroxylation in tumor tissue, and identified specific peptides which are distinctive between CRLM and liver tissue. Comparison with previous data (21) showed that a NOP, which overlaps with peptide 34, was significantly different in urine of patients suffering from CRLM and in urine of healthy controls. This NOP is a promising candidate to improve the detection of CRLM in urine. Especially, because this peptide shows a biologically relation to the liver metastasis tumor. If the degree of collagen hydroxylation is too low, then the T_m of the formed triple helix can be too low and as a result this collagen will not be built into the supramolecular structures. Consequently we may find that there is a possibility to find NOPs in urine that have an even lower degree of hydroxylation than observed in this study, which could possibly be even more distinctive.

Hypoxia is a possible explanation of the overall decreased degree of hydroxylation in tumor tissue. McKeown clarified differences in the partial pressure of oxygen (pO_2) in healthy and tumor tissue (29). The pO_2 in rectal cancer tissue (30 mm Hg) is ∼2 times lower than in healthy rectal tissue (52 mmHg); the pO_2 in liver tumor (6.0 mmHg) is ∼5 times lower than in healthy liver tissue (30 mmHg). Utting et al. (16) showed that hypoxia reduces the degree of hydroxylation and has the largest effect on Hyl formation, and a relative small effect on Hyp formation. In contrast to the above reasoning, the hydroxylation degree of peptides 1 and 21 is decreased in liver tissue in comparison to colon tissue. Peptide 15 and 16, also in contrast to the overall trend, have an increasing hydroxylation degree from CRC to CRLM to liver tissue. Hypoxia can explain the general trend observed, although specific cases such as peptides 1, 15, 16, and 21 do not agree with these observations of McKeown en Utting et al. (16) and may be explained by unknown mechanisms by which hydroxylation is regulated specifically.

Proline and lysine can be enzymatically hydroxylated. Van Huizen et al. (20) described the general upregulation of enzymes related to collagen production and hydroxylation in CRLM. Except the specific enzyme P4HB, the beta unit of 4-prolyl hydroxylase, is strongly reduced in CRLM (11, SD = 5 unique peptides) compared to liver tissue (20, SD = 5 unique peptides) (fold change = 0.48, $p = 1.6^{*}10^{-18}$). P4HB is crucial for the solubility and folding of 4-prolyl hydroxylase. In the present study in colon tissue, on average, 12.8 peptides (SD = 4.1) belonging to P4HB were identified, and in CRC on average 16.1 peptides (SD = 3.9). The downregulation of P4HB in CRLM in comparison to liver tissue might explain the decreased degree of hydroxylation; in CRC tissue, however, P4HB is higher than in colon tissue while the hydroxylation degree is also decreased.

The extracellular matrix is involved in tumor development (30). It has been proposed that the extracellular matrix can act as a tumor suppressor as long as it remains in its'natural shape'; if the "natural morphology" is lost, then it could become a tumor promoter (30). In this study we have shown that besides the collagen expression, also the collagen hydroxylation is significantly different between tumor and normal tissue. Further research is required to understand the effect of the reduced degree

of collagen hydroxylation on tumor proliferation and to identify the mechanisms that are involved in the regulation of changed hydroxylation levels in cancer.

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 at: PRIDE database (PXD015015).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the MERC (medical ethical review committee), Erasmus MC. Written informed consent from the patients was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

NH, PB, JI, and TL designed research. NH, PB, and MD performed research. NH and JR contributed to new analytical tools. NH and PB analyzed data. All authors were involved in writing the paper.

FUNDING

This activity has received funding from the EIT Health under grant agreement no. 19622. This European body receives support from the Horizon 2020 research and innovation program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.557737/full#supplementary-material

REFERENCES

1. Ramshaw JA, Shah NK, Brodsky B. Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *J Struct Biol.* (1998) 122:86–91. doi: 10.1006/jsbi.1998.3977
2. Gorres KL, Raines RT. Prolyl 4-hydroxylase. *Crit Rev Biochem Mol Biol.* (2010) 45:106–24. doi: 10.3109/10409231003627991
3. van Huizen NA, Burgers PC, Saintmont F, Brocorens P, Gerbaux P, Stingl C, et al. Identification of 4-hydroxyproline at the xaa position in collagen by mass spectrometry. *J Proteome Res.* (2019) 18:2045–51. doi: 10.1021/acs.jproteome.8b00930
4. Tianen P, Pasanen A, Sormunen R, Myllyharju J. Characterization of recombinant human prolyl 3-hydroxylase isoenzyme 2, an enzyme modifying the basement membrane collagen IV. *J Biol Chem.* (2008) 283:19432–9. doi: 10.1074/jbc.M802973200
5. Eyre DR, Weis MA, Hudson DM, Wu JJ, Kim L. A novel 3-hydroxyproline (3Hyp)-rich motif marks the triple-helical C terminus of tendon type I collagen. *J Biol Chem.* (2011) 286:7732–6. doi: 10.1074/jbc.C110.195768
6. Jenkins CL, Bretscher LE, Guzei A, Raines RT. Effect of 3-hydroxyproline residues on collagen stability. *J Am Chem Soc.* (2003) 125:6422–7. doi: 10.1021/ja034015j
7. Weis MA, Hudson DM, Kim L, Scott M, Wu JJ, Eyre DR. Location of 3-hydroxyproline residues in collagen types I, II, III, and V/XI implies a role in fibril supramolecular assembly. *J Biol Chem.* (2010) 285:2580–90. doi: 10.1074/jbc.M109.068726
8. Yamauchi M, Sricholpech M. Lysine post-translational modifications of collagen. *Essays Biochem.* (2012) 52:113–33. doi: 10.1042/bse0520113
9. Perdivara I, Yamauchi M, Tomer KB. Molecular characterization of collagen hydroxylysine O-glycosylation by mass spectrometry: current status. *Aust J Chem.* (2013) 66:760–9. doi: 10.1071/CH13174
10. Yamauchi M, Shiiba M. Posttranslational modifications of proteins. In: Kannicht C, editor, *Lysine Hydroxylation and Crosslinking of Collagen*. Totowa, NJ: Humana Press (2002) 277–90.
11. Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humand, flies and worms. *Trends Genet.* (2004) 20:33–43. doi: 10.1016/j.tig.2003.11.004
12. Mizuno K, Buodko SP, Engel J, Bächinger HP. Kinetic hysteresis in collagen folding. *Biophys J.* (2010) 98:3004–14. doi: 10.1016/j.bpj.2010.03.019
13. Leikina E, Mertts MV, Kuznetsova N, Leikin S. Type I collagen is thermally unstable at body temperature. *Proc Natl Acad Sci USA.* (2002) 99:1314–8. doi: 10.1073/pnas.032307099
14. Rapaka RS, Renugopalakrishman V, Urry DW, Bhatnagar RS. Hydroxylation of proline in polytripeptide models of collagen: stereochemistry of polytripeptide-prolyl hydroxylase interaction. *Biochemistry.* (1978) 17:2892–8. doi: 10.1021/bi00607a030
15. Torre-Blanco A, Adachi E, Hojima Y, Wootton JA, Minor RR, Prockop DJ. Temperature-induced post-translational over-modification of type I procollagen. Effects of over-modification of the protein on the rate of cleavage by procollagen N-proteinase and on self-assembly of collagen into fibrils. *J Biol Chem.* (1992) 267:2650–5.
16. Utting JC, Robins SP, Brandao-Burch A, Orriss IR, Behar J, Arnett TR. Hypoxia inhibits the growth, differentiation and bone-forming capacity of rat osteoblasts. *Exp Cell Res.* (2006) 312:1693–702. doi: 10.1016/j.yexcr.2006.02.007
17. Montgomery H, Rustogi N, Hadjisavvas A, Tanaka K, Kyriacou K, Sutton CW. Proteomic profiling of breast tissue collagens and site-specific characterization of hydroxyproline residues of collagen alpha-1-(I). *J Proteome Res.* (2012) 11:5890–902. doi: 10.1021/pr300656r
18. Nissen NI, Karsdal M, Willumsen N. Collagens and cancer associated fibroblasts in the reactive stroma and its relation to cancer biology. *J Exp Clin Cancer Res.* (2019) 38:115. doi: 10.1186/s13046-019-1110-6
19. Naba A, Clauser KR, Whittaker CA, Carr SA, Tanabe KK, Hynes RO. Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC Cancer.* (2014) 14:518. doi: 10.1186/1471-2407-14-518
20. van Huizen NA, Coebergh van den Braak R, Doukas M, Dekker LJ, IJzermans JNM, Luider TM. Up-regulation of collagen proteins in colorectal liver metastasis compared with normal liver tissue. *J Biol Chem.* (2018) 294:281–9. doi: 10.1074/jbc.RA118.005087
21. van Huizen NA, van Rosmalen J, Dekker LJM, Coebergh van den Braak RRJ, Verhoef C, IJzermans JNM, et al. Identification of a collagen marker in urine improves the detection of colorectal liver metastases. *J Proteome Res.* (2020) 19:153–60. doi: 10.1021/acs.jproteome.9b00474
22. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (2016). Retrieved from https://www.R-project.org/ (accessed April 1, 2020).
23. Persikov AV, Ramshaw JA, Brodsky B. Prediction of collagen stability from amino acid sequence. *J Biol Chem.* (2005) 280:19343–9. doi: 10.1074/jbc.M501657200
24. Molony MS, Quan C, Mulkerrin MG, Harris RJ. Hydroxylation of Lys residues reduces their susceptibility to digestion by trypsin and lysyl endopeptidase. *Anal Biochem.* (1998) 258:136–7. doi: 10.1006/abio.1997.2513
25. Siepen JA, Keevil EJ, Knight D, Hubbard SJ. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *J Proteome Res.* (2007) 6:399–408. doi: 10.1021/pr060507u

26. Slechtova T, Gilar M, Kalikova K, Tesarova E. Insight into trypsin miscleavage: comparison of kinetic constants of problematic peptide sequences. *Anal Chem.* (2015) 87:7636–43. doi: 10.1021/acs.analchem.5b00866
27. Broker MEE, Lalmahomed ZS, Roest HP, van Huizen NA, Dekker LJM, Calame W, et al. Collagen peptides in urine: a new promising biomarker for the detection of colorectal liver metastases. *PLoS ONE.* (2013) 8:e70918. doi: 10.1371/journal.pone.0070918
28. Lalmahomed ZS, Broker MEE, van Huizen NA, van den Braak R, Dekker LJM, Rizopoulos D, et al. Hydroxylated collagen peptide in urine as biomarker for detecting colorectal liver metastases. *Am J Cancer Res.* (2016) 6:321–30.
29. McKeown SR. Defining normoxia, physoxia and hypoxia in tumours-implications for treatment response. *Br J Radiol.* (2014) 87:20130676. doi: 10.1259/bjr.20130676
30. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med.* (2011) 17:320–9. doi: 10.1038/nm.2328

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 © 2020 van Huizen, Burgers, van Rosmalen, Doukas, IJzermans and Luider. 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 Oncology

ORIGINAL RESEARCH
published: 28 October 2020
doi: 10.3389/fonc.2020.580891

Proteomic Profiling of Plasma and Total Blood Exosomes in Breast Cancer: A Potential Role in Tumor Progression, Diagnosis, and Prognosis

Oleg Tutanov[1], *Ksenia Proskura*[1,2], *Roman Kamyshinsky*[3,4,5], *Tatiana Shtam*[3,6], *Yuri Tsentalovich*[7] *and Svetlana Tamkovich*[1,8]*

[1] Laboratory of Molecular Medicine, Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russia, [2] Department of Mammology, Novosibirsk Regional Clinical Oncological Dispensary, Novosibirsk, Russia, [3] Probe and Electron Microscopy Resource Centre, National Research Center Kurchatov Institute, Moscow, Russia, [4] Institute of Nano-, Bio-, Information, Cognitive and Socio-humanistic Sciences and Technologies, Moscow Institute of Physics and Technology, Moscow, Russia, [5] Electron Microscopy Laboratory, Shubnikov Institute of Crystallography of Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Moscow, Russia, [6] Laboratory of Protein Biosynthesis, Petersburg Nuclear Physics Institute Named by B.P. Konstantinov of National Research Center "Kurchatov Institute", Gatchina, Russia, [7] Laboratory of Proteomics and Metabolomics, International Tomography Center, Siberian Branch of Russian Academy of Sciences, Novosibirsk, Russia, [8] Department of Molecular Biology and Biotechnology, Novosibirsk State University, Novosibirsk, Russia

OPEN ACCESS

Edited by:
Harsha Gowda,
The University of Queensland, Australia

Reviewed by:
Maria Francesca Baietti,
VIB KU Leuven Center for Cancer Biology, Belgium
Ravi Manoharan,
University of Madras, India

***Correspondence:**
Svetlana Tamkovich
s.tamkovich@g.nsu.ru

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 09 July 2020
Accepted: 16 September 2020
Published: 28 October 2020

Citation:
Tutanov O, Proskura K, Kamyshinsky R, Shtam T, Tsentalovich Y and Tamkovich S (2020) Proteomic Profiling of Plasma and Total Blood Exosomes in Breast Cancer: A Potential Role in Tumor Progression, Diagnosis, and Prognosis. Front. Oncol. 10:580891. doi: 10.3389/fonc.2020.580891

Exosomes are directly involved in governing physiological and pathological processes of an organism by horizontal transfer of functional molecules (proteins, microRNA, etc.) from producing to receiving cells. We explored the relationship of proteins from plasma exosomes, and exosomes from the total blood of healthy females (HFs) and breast cancer patients (BCPs), with crucial steps of tumor progression: EMT, cell proliferation, invasion, cell migration, stimulation of angiogenesis, and immune response. A proteomic analysis of exosomes isolated from samples using ultrafiltration and ultracentrifugation was performed. Their nature has been verified using cryo-electron microscopy and flow cytometry. Bioinformatics analysis showed that 84% of common exosomal proteins were of cytoplasmic and vesicle origin. They perform functions of protein binding and signaling receptor binding, and facilitated the processes of the regulated exocytosis and vesicle-mediated transport. Half of the identified exosomal proteins from blood of HFs and BCPs are involved in crucial steps of the tumor progression: EMT, cell proliferation, invasion, cell migration stimulation of angiogenesis, and immune response. Moreover, we found that protein cargo of exosomes from HF total blood was enriched with proteins inhibiting EMT, cell migration, and invasion. Tumor diagnostic/prognostic protein markers accounted for 47% of the total composition of cell-surface-associated exosomes (calculated as the difference between the total blood exosomes and plasma exosomes) from BCP blood. Breast cancer-associated proteins were equally represented in the blood cell-surface-associated exosomes and in the plasma exosomes from BCPs. However, hyper-expressed proteins predominate in the blood cell-surface-associated exosomes as compared to the plasma exosomes (64 vs. 14%). Using breast cancer

proteins data from the Human Protein Atlas (HPA) (www.proteinatlas.org/), three favorable (SERPINA1, KRT6B, and SOCS3), and one unfavorable (IGF2R) prognostic protein markers were found in the BCP total blood exosomes. Identified exosomal proteins from BCP blood can be recommended for further testing as breast cancer diagnostic/prognostic biomarkers or novel therapeutic targets.

Keywords: exosomes, extracellular vesicles, mass-spectrometry, proteomics, breast cancer, blood

INTRODUCTION

Cancer is a complex disease in which cells with tumorigenic traits (i.e., unregulated cell proliferation and resistance to cell death) have an ability to communicate with neighboring cells to initiate and facilitate the tumor progression. Although malignant cells themselves are the major source of tumorigenesis, their interactions with the tumor microenvironment are critical for the progression from a single tumor mass to distant metastases (1). It was shown that exosomes carry messages (proteins, microRNA, etc.) from transformed to healthy cells or to other cells in the tumor or they may signal in an autocrine manner back to the producing tumor cells causing changes in the recipient cells behavior and microenvironment alterations (2). Notably, exosomes may deliver signals to induce EMT, migration, invasion, angiogenesis, and metastasis processes (2–6). It has been shown blood circulating exosomes contain two fractions of free and cell-surface-associated vesicles (7, 8). Thus, the transport of vesicles is provided by both plasma exosomes and blood cell-surface-associated exosomes. It is known that plasma exosomes and cell-bound exosomes from the blood of cancer patients contain tumor-specific microRNAs (7) and proteins (9). Nonetheless, most studies of the involvement of blood exosomes in carcinogenesis focus on the study of plasma exosomes only. However, another fraction of exosomes—blood cell-surface-associated exosomes—is similarly noteworthy (7, 9).

The goal of the present study is to shed light on the differences between the proteomic cargos of free-circulating and total blood exosomes of healthy females (HFs) and breast cancer patients (BCPs), and to identify exosomal proteins involved in the crucial steps of the tumor progression: EMT, cell proliferation, invasion, cell migration stimulation of angiogenesis, and immune response.

MATERIALS AND METHODS

Ethics Statement

The study protocol was approved by the Ethics Committee of the Institute of Chemical Biology and Fundamental Medicine. Human samples were obtained according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from every female. Blood samples from HFs (n = 21, age 38–65 years, and median age 56) were obtained from Novosibirsk Central Clinical Hospital. HFs did not have any female-related disorders (mastopathy, endometriosis, etc.) or malignant diseases. Blood samples from previously untreated BCPs (n = 23, age 31–78 years, and median age 58) were obtained from Novosibirsk Regional Oncology Dispensary (**Table 1**).

Exosome Isolation

Venous blood (9 mL) was collected in K_3EDTA vacutainers (Improvacuter, China, cat. no. 694091210), immediately mixed using a rotary mixer, placed at +4°C and processed within an hour after blood taking.

The blood sample was divided into two equal parts. One part was used for the isolation of plasma exosomes, and the second—for the isolation of total blood exosomes. To isolate plasma exosomes, blood cells were pelleted by centrifugation at 290 g for 20 min, blood plasma was transferred to a new tube and centrifuged a second time at 1,200 g for 20 min. To remove cell debris, plasma samples were centrifuged at 17,000 g at 4 °C for 20 min.

Plasma supernatants were diluted in PBS (P4417, Sigma, USA) with 5 mM EDTA in a 1:5 ratio, passed through a 100 nm pore-size filter (Minisart high flow, 16553-K, Sartorius, Germany), and the filtrates were centrifuged for 90 min at 100,000 g (4°C). Pellets were suspended in 12 mL PBS and again centrifuged for 90 min at 100,000 g (4°C). This washing stage was performed three times. Then supernatants were removed, and the pellets were resuspended in 150 μL PBS.

TABLE 1 | Clinical characteristics of BCPs.

		No (%)
Tumor stage	T1–T2	20 (87)
	T3–T4	3 (13)
Nodal status	N0	14 (62)
	N1–N2	8 (34)
	N3	1 (4)
	M0	23 (100)
Receptor status	Estrogen and/or Progesterone Positive	15 (65)
	Estrogen and/or Progesterone Negative	5 (22)
	Unknown	3 (13)
HER2	Positive	15 (65)
	Negative	5 (22)
	Unknown	3 (13)
Histologic grade	II	16 (70)
	III	2 (9)
	Unknown	5 (22)
Infiltrative ductal carcinoma		23 (100%)
Total patients		23 (100%)

To isolate total blood exosomes, a previous protocol was used (9). Briefly, equal volumes of elution buffer (BioSilica Ltd., Russia) were added to blood samples, and incubated on a rotary mixer (10 rpm) for 4 min at room temperature. Total blood exosomes were isolated using the same procedures as for plasma exosome isolation.

For cryo-electron microscopy (cryo-EM) study 1/10 of the exosome samples from different individuals were mixed to generate two samples (plasma exosomes and total blood exosomes) from HFs and two samples (plasma exosomes and total blood exosomes) from BCPs. Individual and mixed samples were frozen in liquid nitrogen and stored in aliquots at −80°C until required.

Cryo-Electron Microscopy

Morphology of the isolated exosomes was assessed by cryo-EM using Titan Krios 60–300 TEM/STEM (Thermo Fisher Scientific, USA), equipped with TEM direct electron detector Falcon II (Thermo Fisher Scientific, USA) and Cs image corrector (CEOS, Germany) at accelerating voltage of 300 kV as described (10). The initial volumes of plasma or blood for the study of exosomes using the cryo-EM were 10 μL.

Protein Quantification

To evaluate the protein concentration of exosomes, a NanoOrange Protein Quantitation kit (NanoOrange® Protein Quantitation Kit, Molecular Probes, USA) was used in accordance with the manufacturer's recommendations.

Flow Cytometry Analysis

Flow cytometry was performed on the Cytoflex (Becman Coulter, USA), using CytExpert 2.0 Software. The MFI of stained exosomes was analyzed as described (11) and compared to the isotype control (BD bioscience, USA).

Mass Spectrometry Analysis

Individual exosome samples were separated according to their molecular weight using SDS disc electrophoresis. The gels were stained by Coomassie Brilliant Blue R250 (Sigma, United States). The PAAG fragments containing proteins under study were treated using the modified Rosenfeld method. Briefly, PAAG fragments with proteins were washed from Coomassie R250 and SDS with an aqueous solution containing 50% acetonitrile and 0.1% trifluoroacetic acid. Proteins absorbed in the gel were reduced with 45 mM DTT in 0.2 M ammonium bicarbonate at 60°C for 30 min, followed by protein alkylation with 100 mM iodoacetamide in 0.2 M ammonium bicarbonate, at room temperature for 30 min. The gel fragments were dehydrated in 100% acetonitrile. A 0.2 mM trypsin solution (modified by reductive methylation) (Sigma, T6567, USA), in a mixture with 0.1 M ammonium bicarbonate and 5 μM CaCl2, was added to each gel fragment and incubated for 30 min at room temperature. Then, peptide extraction buffer (60 μL) was added to the gel fragments, and samples were incubated for 16–18 h at 37°C. The peptide fragments of exosomal proteins extracted from the gel, concentrated, and desalted as described (9).

Mass-spectra were registered at the Center of Collective Use "Mass spectrometric investigations" SB RAS on an Ultraflex III MALDI-TOF/TOF mass spectrometer (BrukerDaltonics, Germany) as described (9). Exosomal proteins were identified by searching for appropriate candidates in annotated NCBI and SwissProt databases using Mascot software (Matrix Science Ltd., London, www.matrixscience.com/search_form_select.html).

Analysis of Exosomal Proteome Cargoes

Venn diagram illustrating common proteins in exosome samples from plasma and total blood was drawn using a FunRich version 3.1.3 software package using Vesiclepedia function and current versions of FunRich and Vesiclepedia databases (12). GO functional enrichment analysis of all exosomal fractions was performed using STRING software (www.string-db.org) with the base settings for multiple proteins analysis. The redundancy reduction of the obtained data was conducted using REVIGO software (https://www.irb.hr/) with the allowed similarity of <0.5.

Heatmapping of the enrichment results obtained through STRING was conducted via R (https://www.R-project.org), using packages gplots, dendextend, colorspace.

GO profiling of exosomal proteins involved in the cell migration and motility, immune response, vasculature development, and cell proliferation were performed using the QuickGO annotation terms (lists of obtained proteins were searched against GO terms: cell motility (GO:0048870), cell migration (GO:0016477), negative regulation of cell motility (GO:2000146), immune response (GO:0006955), negative regulation of immune response (GO:0050777), vasculature development (GO:0001944) negative regulation of vasculature development (GO:1901343), cell population proliferation (GO:0008283) and negative regulation of cell population proliferation (GO:0008285) (13). The involvement of exosomal proteins in cancer invasion and EMT were routinely analyzed by searching the PubMed database for relevant publications for each protein.

The relationship of exosomal proteins involved in key stages of tumor development was analyzed using Reactome (www.reactome.org) and NCI-Nature Pathway Interaction Database (www.ndexbio.org) by FunRich (http://www.funrich.org/).

Profiling of the exosomal proteins differently expressed during the development of various malignant diseases was performed using the dbDEPC 3.0 database (14). The search for cancer prognostic proteins in exosomal proteomes was conducted using Human Protein Atlas datasets (http://www.proteinatlas.org/) for breast, renal, thyroid, pancreatic, liver, endometric, head and neck, ovarian, stomach urothelial, cervical, lung and colorectal cancers as well as melanoma, and glioma datasets.

Statistical Analysis

Statistical calculations were performed using Statistica 6.0 software and GraphPad PRISM 5 software (GraphPad Software, La Jolla, CA, USA). All data were expressed either as the median with interquartile ranges or as means with standard errors. To evaluate the differences, the Mann–Whitney *U*-test was performed. $P < 0.05$ were considered statistically significant.

RESULTS

Characterization of Exosomes Isolated From Plasma and Total Blood

The morphology of exosomes from plasma and total blood of HFs and BCPs was examined by cryo-EM. In total, the following sets of images have been analyzed in the present work: 100 images of vesicles from plasma and 116 images of vesicles from total blood of HFs (**Figures 1A,B**); 121 images of vesicles from plasma, and 117 images of vesicles from total blood of BCPs (**Figures 1C,D**). More than 90% of the extracellular vesicles were identified as exosome-like vesicles due to the clear presence of the lipid bilayer/membrane. Most of the vesicles were intact and had round or slightly elongated shapes. Single (**Figures 1A–F**), double (**Figures 1A–D,G,H**), double membrane (**Figures 1I–L**), and multilayer exosome-like vesicles (**Figures 1A–D,I,J**) were found in all samples of isolated vesicles from pooled plasma and pooled total blood of donors and patients. Protein aggregates outside EVs were also visualized. It should be noted that single vesicles were predominantly represented in the HF plasma, while in the total blood vesicles their content decreased (from 91 to 37%) with an increase of double vesicles (from 3 to 20%) and double-membrane vesicles (from 0 to 22%) fractions (**Supplementary Table 1**). The morphology of vesicles isolated from plasma and from the total blood of BCPs was comparable with the prevailing of single vesicles (81 and 70%, respectively; **Supplementary Table 1**).

Exosome-like vesicles from plasma of BCPs were significantly larger compared to HFs (135 $\pm$ 9 and 78 $\pm$ 4 nm, respectively, $p < 0.0001$); however, the exosome-like vesicles from total blood had similar sizes (91 $\pm$ 5 and 92 $\pm$ 5 nm, respectively). It should be noted that the total blood exosomes consist of plasma exosomes and blood cell-surface-associated exosomes. Thus, the obtained data indirectly indicate that in HF blood, the blood-cell-surface associated exosome-like vesicles are larger than in plasma ($p < 0.0322$), and in BCP blood it is vice versa ($p = 0.0007$).

The isolated exosome-like vesicles were also inspected for the presence of exosomal markers (CD9, CD24, CD63, and CD81) using flow cytometry (**Table 2**). A subpopulation of CD9/CD24-positive exosome-like vesicles predominated in both plasma and total blood of HFs and BCPs. No significant differences in the MFI of populations of CD9/CD24-positive, CD9/CD63-positive, and CD9/CD81-positive exosome-like vesicles between plasma and total blood of HFs and BCPs were found.

Collectively, obtained data reveal that the extracellular vesicles isolated from plasma and total blood have all properties of exosomes.

Annotation of Identified Exosomal Proteins

A total of 111 and 146 proteins were identified with high confidence ($P < 0.05$) using mass spectrometry in the blood-circulating exosomes from HFs and BCPs, respectively (**Supplementary Tables 2–5**). Of these, only 15% (34 proteins) were common for both groups (**Figure 2**). The percentage of proteins detected exclusively in the total blood exosomes of HFs and BCPs were 27 and 49%, respectively (**Figure 2**). To determine whether these proteins are associated with exosomes, the protein list was analyzed using Funrich to search against publicly available vesicle protein databases—Vesiclepedia and Exocarta. According to the analysis of the Vesiclepedia database (www.microvesicles.org), of the top 163 proteins associated with vesicles, a total of 80 and 114 proteins were identified in the HF and BCP exosomes, respectively (**Figure 3A**). Thus, about 27% of exosomal proteins are identified in our study for the first time as a part of extracellular vesicles; previously, they were not annotated in this database. Moreover, the analysis of the identified exosomal proteins by Exocarta database (www.exocarta.org) reveals that among 100 proteins that have been previously found to be associated with vesicles, a total of 59 and 72 proteins were identified in the HF and BCP exosomes, respectively. Thus, 55% of proteins identified in exosomes in this study were not previously annotated in the Exocarta database (**Figure 3B**).

Functional enrichment analysis of 32 common exosomal proteins was performed using STRING software (www.string-db.org), and the enriched GO terms for cellular components, molecular functions, and biological processes were highlighted (**Figure 4**). It was shown that 84% of exosomal proteins were of cytoplasmic and vesicle origin (**Figure 4A**). They perform functions of the protein binding and signaling receptor binding (**Figure 4B**), and facilitate processes of the regulated exocytosis and vesicle-mediated transport (**Figure 4C**). This highlights the vesicular nature of the majority of common proteins. The protein-protein interaction was mainly concentrated in the relevance among APOB, APOA1 and APOA4, FGA, TF, HPX, FGG, HP, C3, ALB, AHSG, FGB, TTR, HPR, and CLU (**Figure 4**).

Functional enrichment analysis was performed for all proteins from exosomes from plasma and total blood of HFs and BCPs using STRING software (www.string-db.org). The obtained lists of enriched 98 Cellular Components, 48 Molecular Functions, and 459 Biological Processes GO terms with corresponding p-values were clustered and visualized with heatmaps using R.

The universal terms of cellular components in exosomes included "vesicle" and "endocytic vesicle lumen," "extracellular region part," and "extracellular space." It should be noted that exosomal proteins from HF blood exclusively have only five common GO with lg $p > 5$ ("fibrinogen complex," "supramolecular fiber," "platelet dence granule," "endocytic vesicle membrane," and "plasma membrane"). However, the total GO amount with lg $p > 4$ is 11 in these fractions. Only the exosomes from blood of BCPs were enriched for "cytoplasm," "intracellular part," and "perinuclear region of cytoplasm" (**Figure 5**). Moreover, the Molecular Function GO analysis revealed that the common exosomal proteins are involved in "protein binding," "hemoglobin binding," "cholesterol transporter activity," "identical protein binding," and "lipoprotein particle receptor binding." The "heparin binding," "cell adhesion molecule binding," "peptidase regulator activity," "macromolecular complex binding," "chaperone binding," and "cholesterol binding" were the most common terms for both plasma and total blood of HF exosomes, while "oxygen binding," "oxygen transporter activity," "tetrapyrrole binding,"

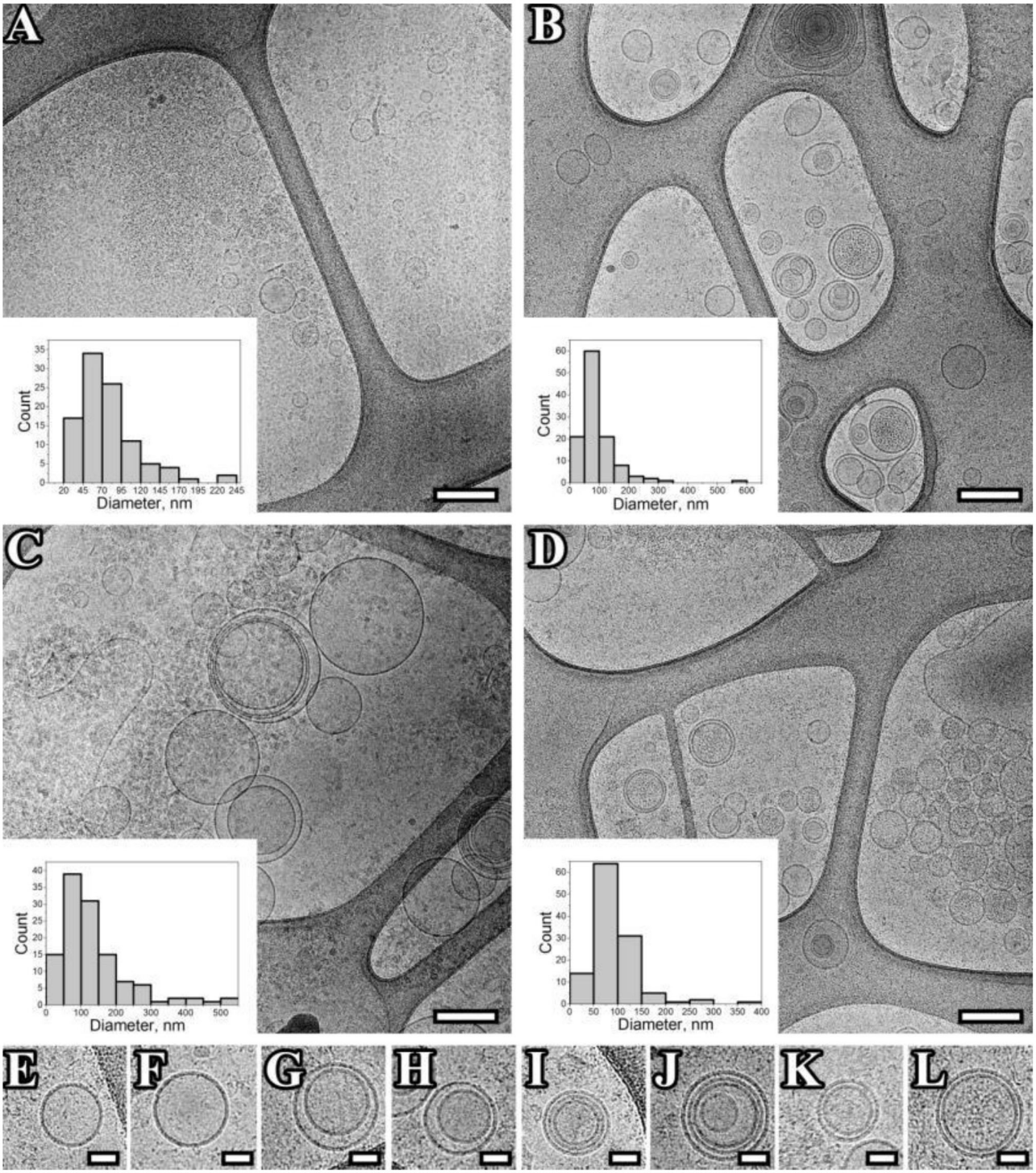


FIGURE 1 | Representation of exosomes with various morphology. **(A–L)**, Cryo-EM images of exosomes isolated from pooled samples of plasma of HFs **(A)**, total blood of HFs **(B)**, plasma of BCPs **(C)**, total blood of BCPs **(D)**, single vesicles **(E,F)**, double vesicles **(G,H)**, multilayer vesicles **(I,J)**, double-membrane vesicles **(I–K)**, vesicles with electron dense cargo in lumen **(G,L)**. Scale bars are 500 nm for micrographs **(A–D)**, and 100 nm for micrographs **(E–L)**.

"microtubule binding," and "small molecule binding" were the most common for plasma and total blood of BCP exosomes (**Figure 6**).

Although the results of the enrichment analysis regarding Biological Processes provided a detailed and broad picture, due to the number of resultant terms it was decided to put

TABLE 2 | Expression of CD 24, CD 63, and CD 81 on the surface of CD9-positive exosomes from plasma and total blood of HFs and BCPs*.

Source of exosomes		CD 24 expression	CD 63 expression	CD 81 expression
HFs	Plasma	1,080 ± 130	540 ± 97	740 ± 81
	Total blood	865 ± 129	570 ± 74	770 ± 85
BCPs	Plasma	1,200 ± 132	480 ± 57	710 ± 92
	Total blood	1,350 ± 189	515 ± 77	745 ± 89

Data represent median fluorescence intensity (MFI) ± SEM.

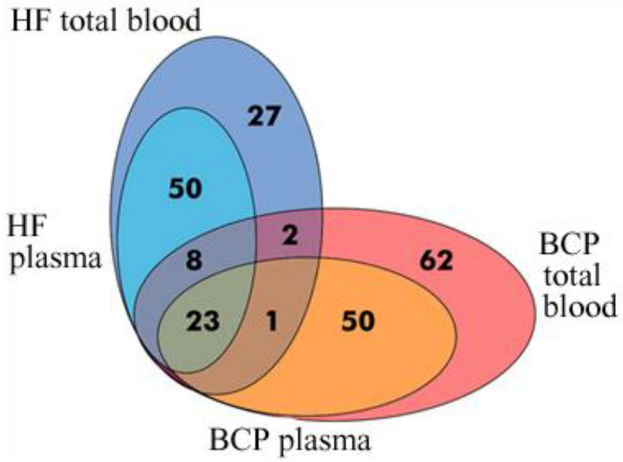


FIGURE 2 | Venn-Euler diagram of proteins in exosomes from HF and BCP blood, composed using FunRich software.

it through the REVIGO software (https://www.irb.hr/) before R visualization to summarize the GO list and to remove the redundant terms (**Figure 7**, **Supplementary Figure 1**). The common proteomic profiles of exosomes from blood of HFs and BCPs are enriched with "endocytosis," "vesicle-mediated transport," "cell activation," "extracellular structure organization," and "regulation of transport" terms. Some terms were enriched exclusively in total blood fractions (that is, only in blood cell-surface-associated exosomes) of both HF and BCP exosomal proteins: "regulation of blood vessel size" and "plasminogen activation." Proteins from BCP exosomes are involved in biological processes such as "response to hypoxia," "protein stabilization," "membrane organization," etc. HF exosomes, on the other hand, were enriched by "regulation of body fluid levels," "protein metabolic process," "transport," "regulation of cell communication," "secretion by cell," etc.

Characterization of Exosomal Proteins Involved in Malignant Neoplasms Development

Literature analysis and QuickGO annotation of the exosomal cargo revealed that many proteins are involved in the crucial steps of the tumor progression: epithelial-to-mesenchymal transition (EMT), cell proliferation, invasion, cell migration, angiogenesis, and immune response (**Table 3**). In particular, a comparable amount of exosomal proteins from total blood of HFs and BCPs (32 (29%) and 35 (24%), respectively; 14 (29%) of which were common to both groups) are involved in EMT. However, the share of the EMT negative regulation proteins in control was two times higher than in disease (22 vs. 11%). Only 14 (13%) proteins from total blood HFs were involved in the cell proliferation, of which several proteins (APOE, CDC73,

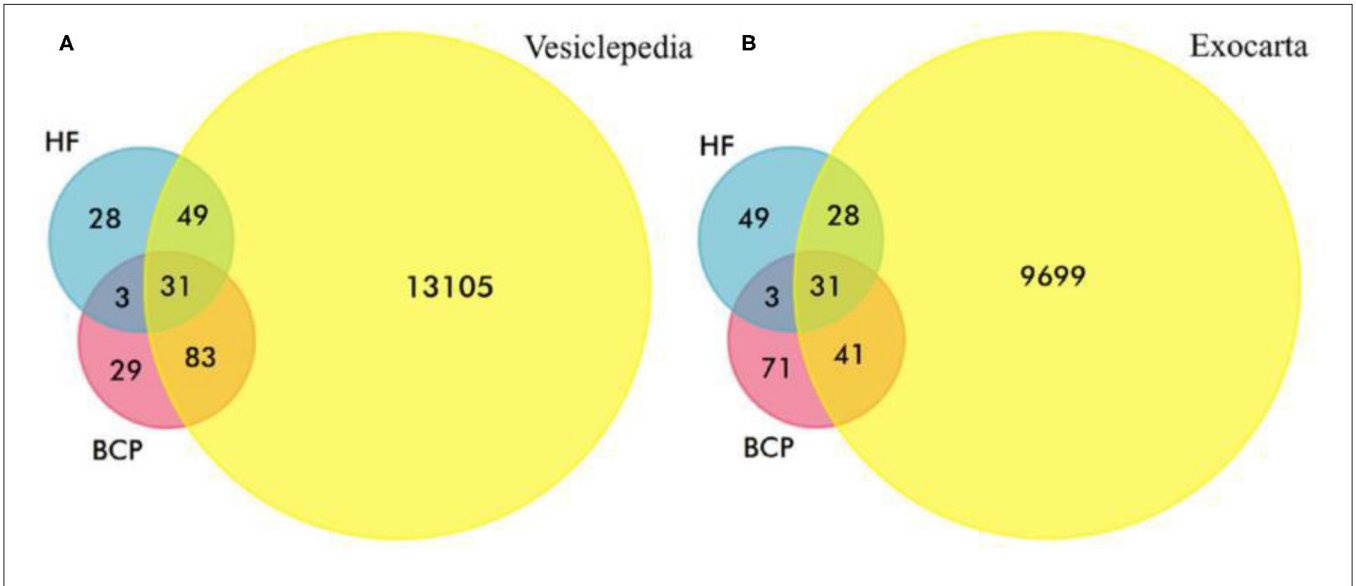


FIGURE 3 | Venn-Euler diagrams of exosomal proteins from HFs and BCPs **(A)** with Vesiclepedia and **(B)** Exocarta databases.

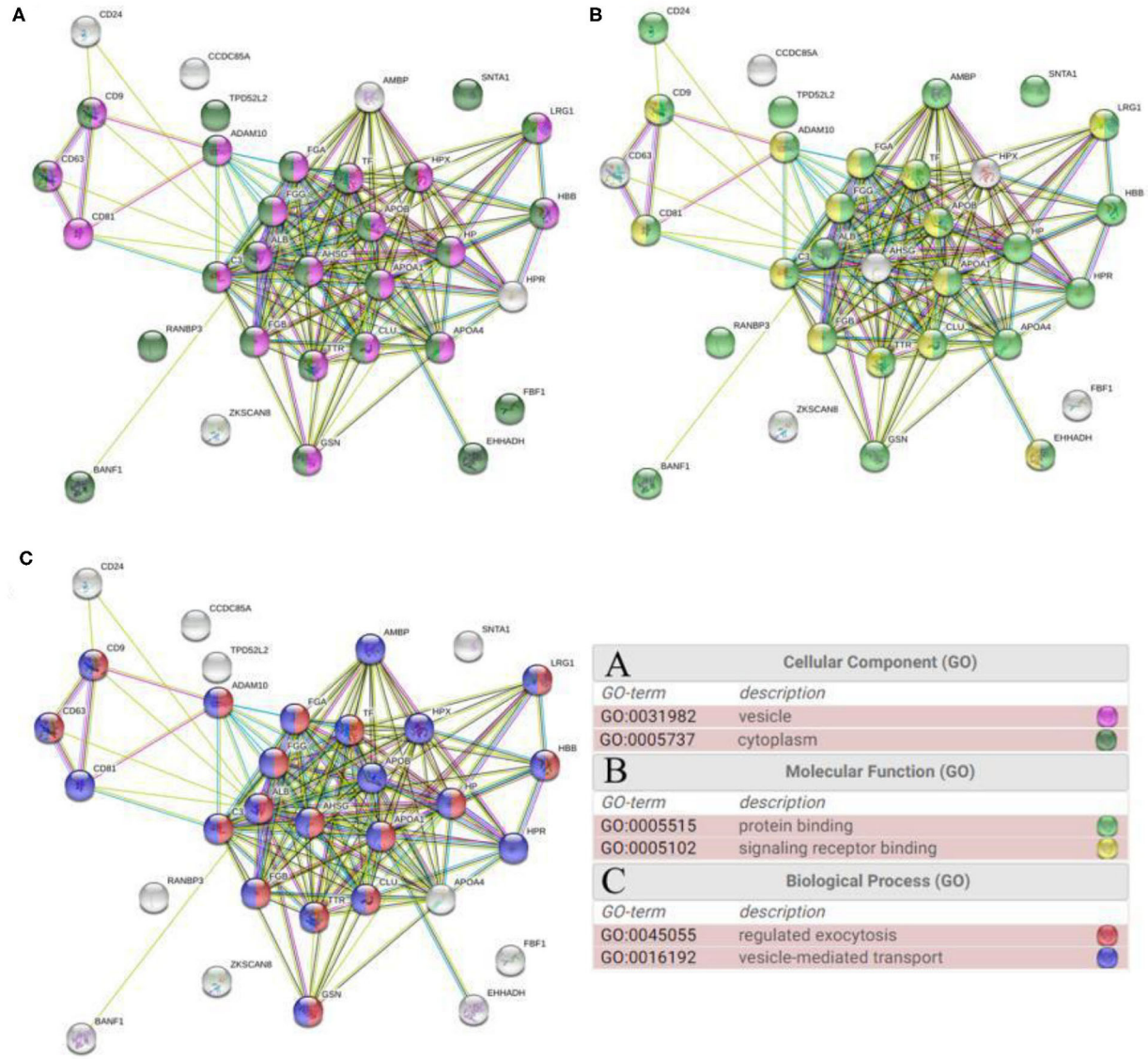


FIGURE 4 | GO functional enrichment analysis of common proteins from total blood exosomes performed using STRING software: **(A)** cellular component, **(B)** molecular function, and **(C)** biological process.

CHD5, FGF10, GDF2, and LMNA) inhibited this process. It is noteworthy that in the BCPs exosome proteome, only 6 (4%) of proteins regulating the cell proliferation were presented, of which four were the process inhibitors. Also, 51 (46%) proteins were involved in the invasion-associated functions in the total blood exosomes from HFs, and 54 (37%) such proteins were revealed in the total BCP blood exosomes, 18 (21%) of which were common for both groups. In the proteome cargo of HF exosomes, 6 (12%) proteins were engaged with the invasion suppression, while only IGF2R and KRT1 (4%) were detected in the exosomes from BCP blood. Moreover, 25 (23%) and 21 (14%) proteins involved in the regulation of the cell migration were revealed in HFs and BCPs total blood exosomes, respectively. GO analysis using QuickGO revealed that four (AKT1, APOE, BARD1, and GDF2) out of 16 unique proteins from exosomes from HF blood were associated with the negative migration regulation of endothelial and lymphocytic cells (15–17). In the vasculature development, 11 (10%) and 6 (4%) proteins of exosomal cargoes from total blood HFs and total blood BCPs were detected, of which only one inhibitor was identified in each group (GDF2 and BMP7, respectively). An equal amount of proteins (30 each) was involved in the regulation of the immune response in the total blood exosomes from HFs and BCPs, the number of the inhibitors in both cases was also equal (3 each) (**Table 3**). Thus, it turned out that the protein cargo of exosomes from HF blood was enriched by proteins for the negative

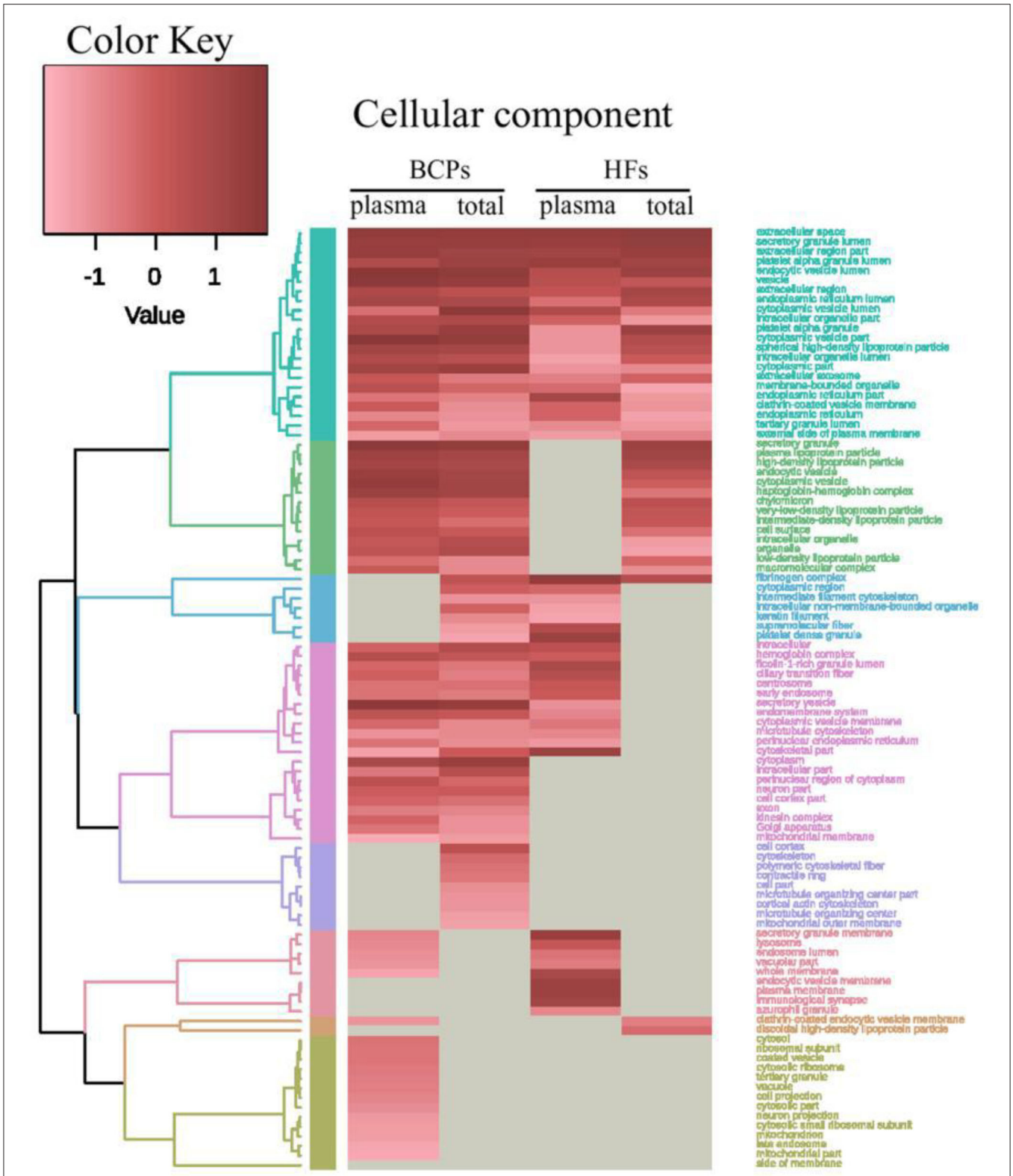


FIGURE 5 | Heat map and dendrogram of the results of GO functional enrichment analysis of exosomal proteins from BCPs and HFs by Cellular Components obtained via STRING (https://string-db.org/) and visualized using R (https://www.r-project.org/).

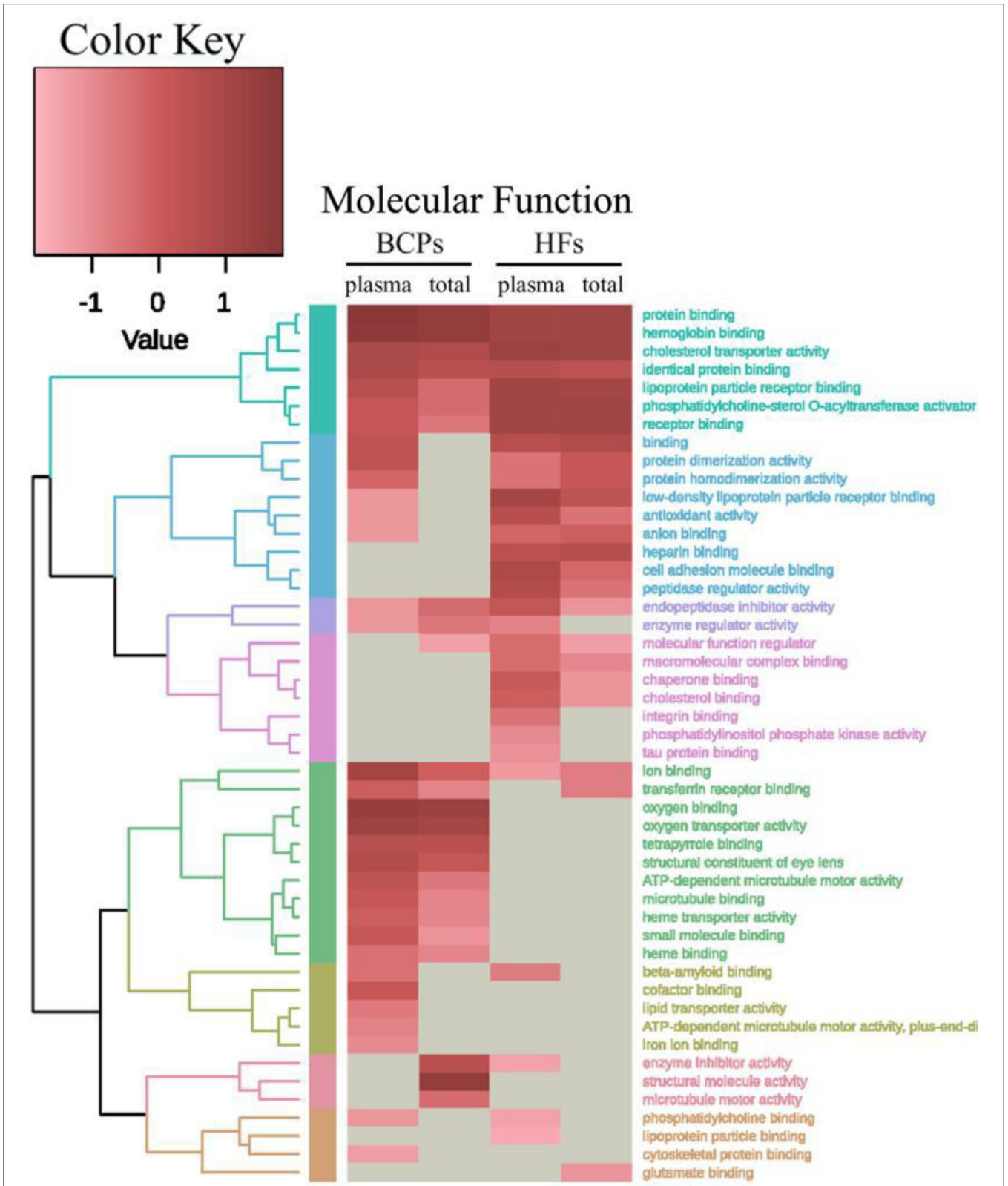


FIGURE 6 | Heat map and dendrogram of the results of GO functional enrichment analysis of exosomal proteins from BCPs and HFs by Molecular Functions obtained via STRING (https://string-db.org/) and visualized using R (https://www.r-project.org/).

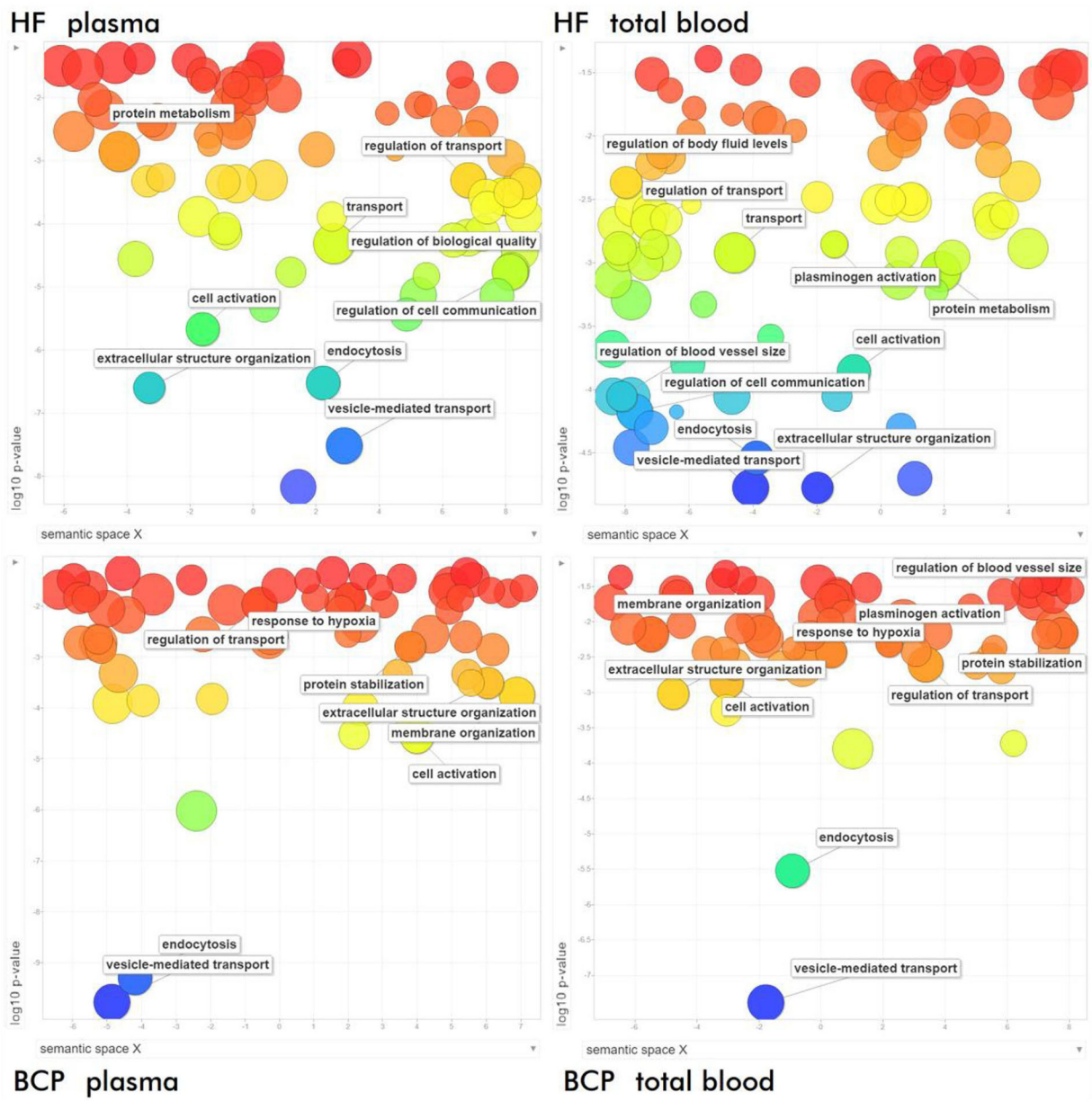


FIGURE 7 | Semantic similarity-based scatterplots of the results of GO functional enrichment analysis of exosomal proteins from BCPs and HFs by Biological Processes obtained via STRING (https://string-db.org/) after redundancy reduction via REVIGO (https://www.irb.hr/).

regulation of EMT, cell proliferation, migration, and invasion. FunRich pathway analysis was used to visualize the interaction network of these proteins. The PPI was mainly concentrated in the relevance among TRAIL signaling pathway (Reactome: R-HSA-75158), Hemostasis (Reactome: R-HSA-109582), and Integrin family cell surface interactions (Reactome: R-HSA-216083) (**Figure 8**).

Using the dbDEPC 3.0 database (database of Differently Expressed Proteins in Human Cancer) (14) it was revealed that proteins associated with the breast cancer development were equally represented in the blood cell-surface-associated exosomes, calculated as the difference between the total blood exosomes and plasma exosomes, and in the plasma exosomes from BCPs [11 (15%) and 14 (19%) proteins, respectively] (**Figure 9**). However, the hyper-expressed proteins predominate in the blood cell-surface-associated exosomes as compared to the plasma exosomes (64 vs. 14%).

Datasets of favorable/unfavorable prognostic proteins for all cancer types from the Human Protein Atlas (HPA) (www.proteinatlas.org/) were used to search for such proteins in the unique exosomal protein cargo of BCPs (**Supplementary Table 6**). In the BCP total blood exosomes, 64 prognostic proteins

TABLE 3 | Exosomal proteins from HFs and BCPs blood associated with crucial steps of tumor progression*.

	Invasion	EMT	Cell Migration	Vasculature development	Immune response	Cell Proliferation
A1BG	BCP				BCP	
A2M	HF				HF*	
ACSM3	BCP					
ADAM10	Universal	Universal	Universal		Universal	
AHSG	Universal	HF*			Universal	
AIFM1		HF*				
AIM2	HF*	HF*			HF	
AKT1	HF	HF*	HF*	HF		HF
AMBP	Universal				HF*	
APOA1	Universal	HF*	Universal		HF*	Universal
APOA4					Universal	
APOB			Universal	Universal		
APOE	HF*		HF*	HF		HF*
APPBP2	BCP					
ATR		BCP				
BANF1	Universal	Universal				
BARD1	HF*	HF	HF*			
BMP1		BCP				
BMP7	BCP	BCP	BCP	BCP*		BCP*
C3	Universal	Universal	Universal		Universal	
CAPN2	BCP	BCP				
CDC73						HF*
CCDC85A	Universal					
CD24	Universal		Universal			
CD63	Universal	Universal	Universal		Universal	
CD81	Universal	Universal	Universal		Universal	
CD9	Universal	Universal	Universal			HF*
CFH					HF	
CHD5	HF	HF				HF*
CKAP2L	BCP					
CKMT2	BCP					
CLK3	BCP					
CLU	Universal				Universal	Universal
COPS2						HF
CORO6			HF			
CPS1	HF					
CRYAB	BCP	BCP				
DKC1	HF					
DNM1L	BCP					
ELL2	BCP					
EXOC8		BCP				
FGA	Universal	Universal			Universal	
FGB		Universal			Universal	
FGF10	HF	HF	HF	HF		HF*
FGG	Universal	Universal				
FN1	HF		HF	HF		
FXYD5	HF					
GCLC	HF	HF				
GDF2	HF*	HF	HF*	HF*		HF*
GRSE1		BCP				
GSN	Universal	Universal			Universal	
GTSE1	BCP		BCP			
HAND1		HF		HF		
HBB					Universal	
HIG1AN	BCP					
HMOX1	BCP	BCP*	BCP	BCP	BCP*	BCP*
HP	Universal	Universal			Universal	
IGF2R	BCP				BCP	
IGHA1			BCP		BCP	
IGHG1	BCP	BCP	BCP		BCP	
IGHG2					Universal	
IGHM			HF		HF	
IGHV3-74					Universal	
IGKC			Universal		Universal	
IGKV3-20			BCP		BCP	
IGLC7			HF		HF	
IL16	BCP	BCP	BCP		BCP	
ITIH4	BCP		BCP			

(Continued)

TABLE 3 | Continued

	Invasion	EMT	Cell Migration	Vasculature development	Immune response	Cell Proliferation
KDM6B	HF	HF		HF	HF	
KIF20B	BCP	BCP	BCP			
KIF3B	BCP					
KIFC3	BCP					
KRT1	BCP				BCP	
KRT6A	BCP				BCP	
KRT6B	BCP					
LMNA	HF		HF	HF		HF*
LRG	Universal	Universal			Universal	
MBD4	BCP					
MFAP5	HF	HF				
MIA	HF					
MTHFD1				BCP		
NANOS3	HF	HF				
NAP1L1	HF					
P0DOX6					HF	
PCNT	BCP					
PDZD2	BCP	BCP	BCP			
PGF	HF	HF	HF	HF		
PHB2	BCP					HF*
PIBF1	BCP					
PIK3CD	HF	HF	HF		HF	
PIP4K2B	HF*	HF*				
PLEKHA7	HF*					
PPM1A	BCP	BCP*				
PSKH1	HF		HF	HF		
RAB22A	HF	HF*				
RAB24		BCP			BCP	
RANBP3	HF					
RAP1B	HF				HF	HF
RASGRP4						HF
RASSF2	HF					
RFC3	HF	HF				
SERPINA1	BCP				BCP	
SERPINB7	BCP					
SOCS3	BCP	BCP		BCP		
ST3GAL1	HF	HF	HF			
TNFSF14	HF		HF		HF	HF
TPD52L2	Universal					
TTR	HF				Universal	
UBA52	BCP					
VAV3	BCP	BCP	BCP	BCP		
VPS13A	BCP	BCP	BCP			
ZAP70	HF		HF		HF	

**Proteins, that negatively regulate the process are marked with hatching.*

HF BCP Universal

typical for different types of cancer were detected, and only 29 in the plasma exosomes. Thus, 35 of these 64 prognostic proteins found in the exosomes from total blood were detected solely through the analysis of exosomes associated with the surface of blood cells. Using the breast cancer proteomic data from HPA, we revealed three favorable (SERPINA1, KRT6B, and SOCS3) and one unfavorable (IGF2R) tumor prognostic protein markers. Notably, KRT6B and SOCS3 were revealed in the total blood exosomes from BCPs exclusively (i.e., in the composition of the blood cell-surface-associated exosomes). ROC analysis of revealed tumor markers (SERPINA1, KRT6B, SOCS3, and IGF2R) in total blood exosomes demonstrated that combination of these proteins enabled to distinguish BCPs at stage I and II of the disease with sensitivity of 73% and specificity of 100%

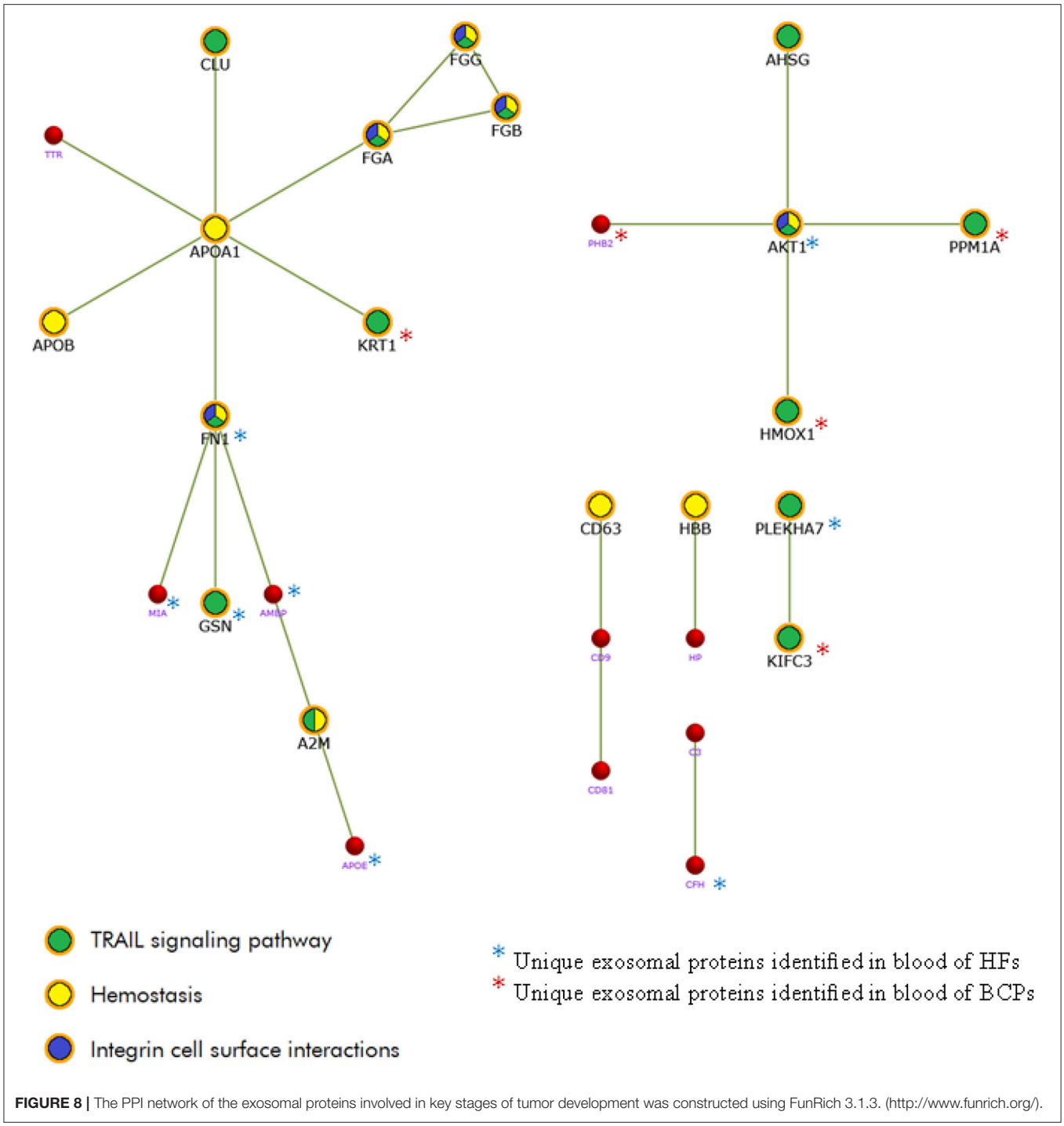


FIGURE 8 | The PPI network of the exosomal proteins involved in key stages of tumor development was constructed using FunRich 3.1.3. (http://www.funrich.org/).

(AUG 0.864), whereas the same analysis in the fraction of cell-free exosomes enabled to distinguish cancer individuals only with 45% sensitivity and 100% specificity.

DISCUSSION

Exosomes are endosome-derived vesicles (30-120 nm) that are ubiquitous in biological fluids [tear (18), urine (19), cerebrospinal fluid (10), ascites (20), blood (21), and saliva (22)]. These vesicles are sufficiently small to penetrate from various tissues into biological fluids and then back, interacting with target cells, or tissues. There is a considerable research interest in understanding the physiological functions of exosomes, and in elucidating their role in diseases like cancer. Most studies are devoted to the study of the exosome proteome from cell cultures (1, 3, 23–27), which may be associated with various

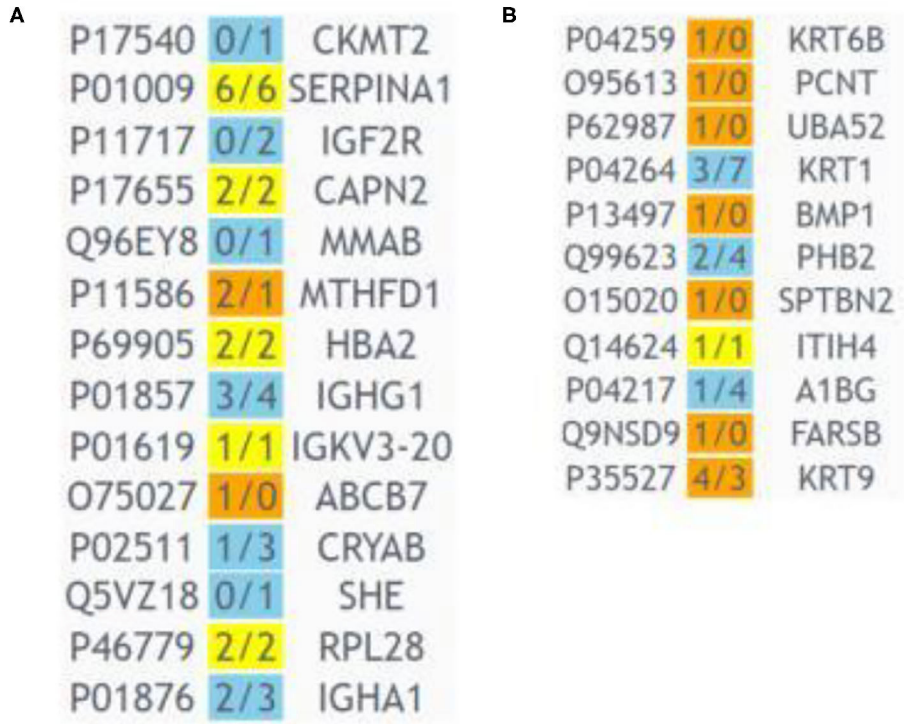


FIGURE 9 | Heat map generated by dbDEPC 3.0 (14) from exosomal proteins from BCP blood, associated with breast cancer. **(A)** Proteins from plasma exosomes, **(B)** proteins from blood cell-surface-associated exosomes. Orange means that the number of the studies identified this protein as up-regulated is more than the number of the studies identified the protein as down-regulated; blue means that the number of the studies identified this protein as up-regulated is less than the number of the studies identified protein as down-regulated; yellow means that the number of the studies identified the protein as up-regulated is equal to the number of the studies identified protein as down-regulated.

technical difficulties in analyzing exosomes from biological fluids (additional purification from concomitant protein impurities, heterogeneity of vesicles, etc.). Nevertheless, the study of exosomes isolated from clinical samples has greater significance and closer association with the real-life situation of patients. In the present, the studies of exosomes from plasma/serum in ovarian (21), breast (24), and colorectal cancers (28) have begun. A recent study has found that roughly two-thirds of the total blood extracellular vesicles of breast cancer patients exists in the blood cell-bound state (7, 8). A thorough investigation of blood cell-bound EVs as significant players of cell-to-cell communication and as potential cancer biomarkers is therefore entirely justified.

Our previously published data indicated that total blood exosomes morphologically resemble plasma exosomes, and that both pools of vesicles are positive for exosomal (CD24, CD63, CD81, and CD9) markers (9). We established with the use of cryo-EM a wide spectrum of the exosome morphologies with lipid bilayers and vesicular internal structures. To our knowledge, only two studies have examined several novel morphological EVs subcategories from human blood plasma by cryo-EM (29, 30) and none from human total blood. We revealed a variety of exosomes (single, double, double with two membrane bilayers and multilayer vesicles), which is in accordance with the previously obtained data (29, 30). Double and multilayered vesicles containing the electron-dense material were also visualized. These types of EVs were described earlier in the ejaculate and cerebrospinal fluid (10, 31). For the first time, we have found that single exosomes predominate in the plasma of HFs and the plasma and total blood of BCPs, while their proportion in the total blood of HFs is reduced by almost 2.5 times. Since the total blood exosomes contain two fractions (cell-free exosomes and blood cell-surface-associated exosomes), the revealed phenomenon indicates that exosomes of various morphologies predominate on the surface of blood cells. It allows suggesting that the subpopulations of exosomes with different and specific functions exist, and blood cells act as their transporters. Moreover, it seems that cancer can cause a redistribution of these subpopulations.

At the next step, using MALDI-TOF mass spectrometry we identified 223 proteins from plasma exosomes and total blood exosomes (containing plasma exosomes and blood-cell-surface associated exosomes) of HFs and BCPs. 27% of the identified proteins were absent in the Vesiclepedia database and 55% in the Exocarta database. This can be attributed to the difference in the sizes of the databases (13,105 Vesiclepedia proteins vs. 9,699 annotated in Exocarta), to the significant heterogeneity of extracellular vesicles, and to the presence of cell-bound exosomal

proteins that elude researchers when only plasma exosomes are studied.

In according to previous studies, our obtained data shows that the main part of detected common exosomal proteins was of cytoplasmic and vesicle origin (3, 21, 28). The assignment of the subcellular localization revealed that a large proportion of unique exosomal proteins from plasma and total blood of BCPs is normally located in the cytoplasm of cells or vesicles; however, in the blood of HFs the unique exosomal proteins with GO term "fibrinogen complex" predominate. Unexpectedly, this GO term was also identified in a proteomic study of plasma exosomes in patients with prostate cancer (32). This finding is consistent the theory that biogenesis and selection of exosomal content is not a random procedure, but rather the result of a selective sorting process. Multilevel GO enrichment analysis of molecular functions revealed that the common exosomal proteins from plasma and total blood are involved in the protein binding, which is consistent with the earlier studies (21, 28, 33). Moreover, the presence of proteins associated with the receptor binding and cell adhesion molecule binding in the proteome cargo of exosomes from HF blood is consistent with the increased content of exosomes on the blood cell surface of healthy donors (7), while the share of blood-cell-associated exosomes is significantly reduced in tumor patients. The presence of proteins associated with the oxygen binding and oxygen transporter activity in the exosome composition is possibly associated with the adaptation of the tumor under conditions of hypoxia (34). The proteins involved in biological processes such as "endocytosis," "vesicle-mediated transport," "cell activation," "extracellular structure organization," and "regulation of transport" are common for the proteomic profiles of exosomes from blood of HFs and BCPs and are consistent with the previous studies (25, 35). An interesting finding was that the protein cargo from BCP exosomes is involved in the response to hypoxia; from the other hand, HF exosomes are enriched by proteins involved in the protein metabolic process, transport, regulation of cell communication, and secretion by cell (21, 28). Thus, the identified proteins from the total blood exosomes may play multiple roles in biological cell functions through the crucial mechanisms. These mechanisms include: (a) contact between surface molecules of exosomes and cells; (b) endocytosis and exocytosis of exosomes; (c) fusion of exosomes-cell membrane. It is possible that the production of exosomes allows neoplastic cells to exert various effects in accordance with possible acceptor targets. For example, exosomes can enhance the malignant properties of neighboring neoplastic cells and/or activate non-malignant adjacent cells through the transfer of multifunctional proteins, which play a key role in the cancerogenesis. To elucidate the molecular regulatory roles underlying blood exosome-mediated tumor progression, an analysis of protein cargoes in plasma and total blood exosomes was performed. Currently, there are few reports on the proteomic analysis of exosomes, especially for plasma exosomes in breast cancer in comparison with other types of cancer (24, 36) and a dearth of data on the proteomic profiles of blood cell-bound exosomes. Technical difficulties in the blood exosome proteomics are primarily associated with the heterogeneity of the composition of circulating vesicles. Nevertheless, we found that half of the identified exosomal proteins are involved in EMT, cell proliferation, invasion, cell migration, vasculature development, and immune response. Moreover, we revealed that the proteomic profiles of exosomes from HF total blood were enriched with the proteins inhibiting EMT, cell proliferation, migration, and invasion. Thus, the analysis of the proteomic composition of exosomes complements their biogenesis and functional roles in the cancer development.

Recently, liquid biopsy has gained much attention in cancer research, since it offers multiple advantages in the clinical settings, including their noninvasive nature, a suitable sample source for longitudinal disease monitoring, a better screenshot of tumor heterogeneity, and so on. Current liquid biopsies primarily focus on the detection and downstream analysis of circulating tumor cells, circulating tumor DNA, and EVs. The disadvantages of circulating tumor cells and circulating tumor DNA are heterogeneity and extreme rarity, while exosomes do not have such problems. Most of the studies on EVs focus on microRNAs and a small portion of exosomal proteins. Detection of functional proteins will provide real-time information about the body's physiological functions and disease progression, such as early detection and monitoring of cancer. In our study, we found that the breast cancer-associated proteins were equally represented in the blood cell-surface-associated exosomes (calculated as the difference between the total blood exosomes and plasma exosomes) and in the plasma exosomes from BCPs. However, the hyper-expressed proteins predominate in the blood cell-surface-associated exosomes as compared to the plasma exosomes (64 vs. 14%). Nevertheless, the use of the total blood exosomes without division into subfractions (plasma exosomes and blood cell-surface-associated exosomes) can reduce the time and labor required for the sample preparation before analysis, as well as increase the amount of the analyzed material, which is an extremely important aspect for further application of the developed methods to clinical practice. Using the data on favorable/unfavorable prognostic proteins in breast cancer from the Human Protein Atlas (HPA) (www.proteinatlas.org/), we found three favorable (SERPINA1, KRT6B, and SOCS3) and one unfavorable (IGF2R) prognostic protein markers in the total blood exosomes. Previously, these proteins were recognized as the prognostic markers (37, 38) in Human Pathology Atlas, and their expression in Breast tissue is linked to the overall higher/lower survival rates with significant ($p < 0.001$) association. Serpin family A member 1 (SERPINA1) is an inhibitor of serine proteases, namely elastase, trypsin, chymotrypsin, and plasminogen activator, which is also secreted to blood; it is involved in the invasion (39) and immune response, was found in vesicles, and it is a prognostic marker for the colorectal cancer (favorable prognosis). Keratin 6B (KRT6B) is a cytoskeleton protein involved in invasion (40) and prognostic in the urothelial cancer (unfavorable). Suppressor of the cytokine signaling 3 (SOCS3) is a suppressor of the cytokine signal transduction, localized in the cytosol and plasma membrane, involved in EMT (41), invasion (42), and vasculature development and prognostic for the renal cancer (unfavorable). Insulin-like growth factor 2 receptor (IGF2R) is a protein that transports phosphorylated lysosomal enzymes from the Golgi

complex and the cell surface to lysosomes. It is situated in the Golgi apparatus and vesicles and acts as a positive regulator of the T-cell coactivation, shown to facilitate the immune response and invasion (43). It is also a prognostic marker for the cervical and urothelial cancers (unfavorable).

Thus, this study demonstrates that the proteomic analysis of exosomes from the total blood of cancer patients reveals more tumor markers. However, to facilitate the clinical use of such biomarker panels, a simple and reliable method for isolating exosomes from total blood should be developed.

In summary, the results of the present study indicate that the protein cargoes of both exosomes from plasma and blood cell-associated exosomes (calculated as the difference between the total blood exosomes and plasma exosomes) of HFs and cancer patients are enriched with proteins involving into the crucial steps of the tumor progression: EMT, cell proliferation, invasion, cell migration, stimulation of angiogenesis, and immune response. We found that the percentage of the tumor diagnostic/prognostic protein markers in the cell-surface-associated exosomes from BCP blood reaches 47%, with over-expressed proteins prevailing in this fraction (64 vs. 14% in the plasma exosomes). Further studies will be needed to determine whether identified new exosomal proteins from BCP blood are potential breast cancer diagnostic/prognostic markers or novel therapeutic targets.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committees of the Novosibirsk Regional Clinical Oncological Dispensary and the Institute of Chemical Biology and Fundamental Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ST: conceptualization. OT, KP, RK, TS, and ST: investigation. OT, RK, TS, YT, and ST: writing. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.580891/full#supplementary-material

REFERENCES

1. Sharma S, Alharbi M, Kobayashi M, Lai A, Guanzon D, Zuniga F, et al. Proteomic analysis of exosomes reveals an association between cell invasiveness and exosomal bioactivity on endothelial and mesenchymal cell migration *in vitro*. *Clin Sci (Lond).* (2018) 132:2029–44. doi: 10.1042/CS20180425
2. Conigliaro A, Cicchini C. Exosome-mediated signaling in epithelial to mesenchymal transition and tumor progression. *J Clin Med.* (2018) 8:26. doi: 10.3390/jcm8010026
3. Harris DA, Patel SH, Gucek M, Hendrix A, Westbroek W, Taraska JW. Exosomes released from breast cancer carcinomas stimulate cell movement. *PLoS One.* (2015) 10:e0117495. doi: 10.1371/journal.pone.0117495
4. Shtam T, Naryzhny S, Samsonov R, Karasik D, Mizgirev I, Kopylov A, et al. Plasma exosomes stimulate breast cancer metastasis through surface interactions and activation of FAK signaling. *Breast Cancer Res Treat.* (2019) 174:129–41. doi: 10.1007/s10549-018-5043-0
5. Li R, Wang Y, Zhang X, Feng M, Ma J, Li J, et al. Exosome-mediated secretion of LOXL4 promotes hepatocellular carcinoma cell invasion and metastasis. *Mol Cancer.* (2019) 18:e18. doi: 10.1186/s12943-019-0948-8
6. Li Z, Tao Y, Wang X, Jiang P, Li J, Peng M, et al. Tumor-secreted exosomal miR-222 promotes tumor progression via regulating P27 expression and re-localization in pancreatic cancer. *Cell Physiol Biochem.* (2018) 51:610–29. doi: 10.1159/000495281
7. Tamkovich S, Tutanov O, Efimenko A, Grigor'eva A, Ryabchikova E, Kirushina N, et al. Blood circulating exosomes contain distinguishable fractions of free and cell-surface-associated vesicles. *Curr Mol Med.* (2019) 19:273–85. doi: 10.2174/1566524019666190314120532
8. Mamaeva SN, Kononova IV, Ruzhansky M, Nikiforov PV, Nikolaeva NA, Pavlov AN, et al. Using scanning electron microscopy and atomic force microscopy to study the formation of nanoparticles on red blood cell surface in cervical cancer patients. *Int J Biomed.* (2020) 10:70–5. doi: 10.21103/Article10(1)_OA12
9. Tutanov O, Orlova E, Proskura K, Grigor'eva A, Yunusova N, Tsentalovich Y, et al. Proteomic analysis of blood exosomes from healthy females and breast cancer patients reveals an association between different exosomal bioactivity on non-tumorigenic epithelial cell and breast cancer cell migration *in vitro*. *Biomolecules.* (2020) 10:495. doi: 10.3390/biom10040495
10. Emelyanov A, Shtam T, Kamyshinsky R, Garaeva L, Verlov N, Miliukhina I, et al. Cryo-electron microscopy of extracellular vesicles from cerebrospinal fluid. *PLoS One.* (2020) 15:e0227949. doi: 10.1371/journal.pone.0227949
11. Tamkovich SN, Yunusova NV, Tugutova E, Somov AK, Proskura KV, Kolomiets LA, et al. Protease cargo in circulating exosomes of breast cancer and ovarian cancer patients. *Asian Pac J Cancer Prev.* (2019) 20:255–62. doi: 10.31557/APJCP.2019.20.1.255
12. Pathan M, Keerthikumar S, Chisanga D, Alessandro R, Ang CS, Askenase P, et al. A novel community driven software for functional enrichment analysis of extracellular vesicles data. *J Extracell Vesicles.* (2017) 6:e1321455. doi: 10.1080/20013078.2017.1321455
13. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics.* (2009) 25:3045–6. doi: 10.1093/bioinformatics/btp536
14. Yang Q, Zhang Y, Cui H, Chen L, Zhao Y, Lin Y, et al. dbDEPC 3.0: The database of differentially expressed proteins in human cancer with multi-level annotation and drug indication. *Database (Oxford).* (2018) 2018:bay015. doi: 10.1093/database/bay015
15. Pencheva N, Tran H, Buss C, Huh D, Drobnjak M, Busam K, et al. Convergent multi-miRNA targeting of ApoE drives LRP1/LRP8-dependent melanoma metastasis and angiogenesis. *Cell.* (2012) 151:1068–82. doi: 10.1016/j.cell.2012.10.028
16. Luo C, Wang F, Ren X, Ke T, Xu C, Tang B, et al. Identification of a molecular signaling gene-gene regulatory network between GWAS susceptibility genes ADTRP and MIA3/TANGO1 for coronary artery disease. *Biochim Biophys Acta Mol Basis Dis.* (2017) 1863:1640–53. doi: 10.1016/j.bbadis.2017.03.010
17. Suzuki Y, Ohga N, Morishita Y, Hida K, Miyazono K, Watabe T. BMP-9 induces proliferation of multiple types of endothelial cells *in vitro* and *in vivo*. *J Cell Sci.* (2010) 123(Pt 10):1684–92. doi: 10.1242/jcs.061556
18. Tamkovich S, Grigor'eva A, Eremina A, Tupikin A, Kabilov M, Chernykh V, et al. What information can be obtained from the tears of a patient

with primary open angle glaucoma? *Clin Chim Acta.* (2019) 495:529–37. doi: 10.1016/j.cca.2019.05.028
19. Overbye A, Skotland T, Koehler CJ, Thiede B, Seierstad T, Berge V, et al. Identification of prostate cancer biomarkers in urinary exosomes. *Oncotarget.* (2015) 6:30357–76. doi: 10.18632/oncotarget.4851
20. Yunusova NV, Patysheva MR, Molchanov SV, Zambalova EA, Grigor'eva AE, Kolomiets LA, et al. Metalloproteinases at the surface of small extrcellular vesicles in advanced ovarian cancer: relationships with ascites volume and peritoneal canceromatosis index. *Clin Chim Acta.* (2019) 494:116–22. doi: 10.1016/j.cca.2019.03.1621
21. Zhang W, Ou X, Wu X. Proteomics profiling of plasma exosomes in epithelial ovarian cancer: a potential role in the coagulation cascade, diagnosis and prognosis. *Int J Oncol.* (2019) 54:1719–33. doi: 10.3892/ijo.2019.4742
22. Xiao H, Wong DT. Proteomic analysis of microvesicles in human saliva by gel electrophoresis with liquid chromatography-mass spectrometry. *Anal Chim Acta.* (2012) 723:61–7. doi: 10.1016/j.aca.2012.02.018
23. Lazar I, Clement E, Ducoux-Petit M, Denat L, Soldan V, Dauvillier S, et al. Proteome characterization of melanoma exosomes reveals a specific signature for metastatic cell lines. *Pigment Cell Melanoma Res.* (2015) 28:464–75. doi: 10.1111/pcmr.12380
24. Vardaki I, Ceder S, Rutishauser D, Baltatzis G, Foukakis T, Panaretakis T. Periostin is identified as a putative metastatic marker in breast cancer-derived exosomes. *Oncotarget.* (2016) 7:74966–78. doi: 10.18632/oncotarget.11663
25. Wang S, Chen G, Lin X, Xing X, Cai Z, Liu X, et al. Role of exosomes in hepatocellular carcinoma cell mobility alteration. *Oncol Lett.* (2017) 14:8122–31. doi: 10.3892/ol.2017.7257
26. Gangoda L, Liem M, Ang CS, Keerthikumar S, Adda CG, Parker BS, et al. Proteomic profiling of exosomes secreted by breast cancer cells with varying metastatic potential. *Proteomics.* (2017) 17:1600370. doi: 10.1002/pmic.201600370
27. Wen SW, Lima LG, Lobb RJ, Norris EL, Hastie ML, Krumeich S, et al. Breast cancer-derived exosomes reflect the cell-of-origin phenotype. *Proteomics.* (2019) 19:e1800180. doi: 10.1002/pmic.201800180
28. Chen Y, Xie Y, Xu L, Zhan S, Xiao Y, Gao Y, et al. Protein content and functional characteristics of serum-purified exosomes from patients with colorectal cancer revealed by quantitative proteomics. *Int J Cancer.* (2017) 140:900–13. doi: 10.1002/ijc.30496
29. Arraud N, Linares R, Tan S, Gounou C, Pasquet J-M, Mornet S, et al. Extracellular vesicles from blood plasma: determination of their morphology, size, phenotype and concentration. *J Thromb Haemost.* (2014) 12:614–27. doi: 10.1111/jth.12554
30. Yuana Y, Koning RI, Kuil ME, Rensen PCN, Koster AJ, Bertina RM, et al. Cryo-electron microscopy of extracellular vesicles in fresh plasma. *J Extracell Vesicles.* (2013) 2:10.3402/jev.v2i0.21494. doi: 10.3402/jev.v2i0.21494
31. Höög JL, Lötvall J. Diversity of extracellular vesicles in human ejaculates revealed by cryo-electron microscopy. *J Extracell Vesicles.* (2015) 4:28680. doi: 10.3402/jev.v4.28680
32. Welton JL, Brennan P, Gurney M, Webber JP, Spary LK, Carton DG, et al. Proteomics analysis of vesicles isolated from plasma and urine of prostate cancer patients using a multiplex, aptamer-based protein array. *J Extracell Vesicles.* (2016) 5:31209. doi: 10.3402/jev.v5.31209
33. Liang B, Peng P, Chen S, Li L, Zhang M, Cao D, et al. Characterization and proteomic analysis of ovarian cancer-derived exosomes. *J Proteomics.* (2013) 80:171–82. doi: 10.1016/j.jprot.2012.12.029
34. Kumar A, Deep G. Exosomes in hypoxia-induced remodeling of the tumor microenvironment. *Cancer Lett.* (2020) 488:1–8. doi: 10.1016/j.canlet.2020.05.018
35. Ge M, Ke R, Cai T, Yang J, Mu X. Identification and proteomic analysis of osteoblast-derived exosomes. *Biochem Biophys Res Commun.* (2015) 467:27–32. doi: 10.1016/j.bbrc.2015.09.135
36. Chen IH, Xue L, Hsu CC, Paez JSP, Pan L, Andaluz H, et al. Phosphoproteins in extracellular vesicles as candidate markers for breast cancer. *Proc Natl Acad Sci U S A.* (2017) 114:3175–80. doi: 10.1073/pnas.1618088114
37. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science.* (2017) 357:eaan2507. doi: 10.1126/science.aan2507
38. Pontén F, Schwenk JM, Asplund A, Edqvist PH. The Human Protein Atlas as a proteomic resource for biomarker discovery. *J Intern Med.* (2011) 270:428–46. doi: 10.1111/j.1365-2796.2011.02427.x
39. Kwon CH, Park HJ, Choi JH, Lee JR, Kim HK, Jo HJ, et al. Snail and serpinA1 promote tumor progression and predict prognosis in colorectal cancer. *Oncotarget.* (2015) 6:20312–26. doi: 10.18632/oncotarget.3964
40. Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, et al. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol.* (2006) 19:264–71. doi: 10.1038/modpathol.3800528
41. Tang H, Long Q, Zhuang K, Yan Y, Han K, Guo H, et al. miR-665 promotes the progression of gastric adenocarcinoma via elevating FAK activation through targeting SOCS3 and is negatively regulated by lncRNA MEG3. *J Cell Physiol.* (2020) 235:4709–19. doi: 10.1002/jcp.29349
42. Chu Q, Shen D, He L, Wang H, Liu C, Zhang W. Prognostic significance of SOCS3 and its biological function in colorectal cancer. *Gene.* (2017) 627:114–22. doi: 10.1016/j.gene.2017.06.013
43. Tian Z, Yao G, Song H, Zhou Y, Geng J. IGF2R expression is associated with the chemotherapy response and prognosis of patients with advanced NSCLC. *Cell Physiol Biochem.* (2014) 34:1578–88. doi: 10.1159/000366361

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 © 2020 Tutanov, Proskura, Kamyshinsky, Shtam, Tsentalovich and Tamkovich. 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 Oncology

ORIGINAL RESEARCH
published: 20 November 2020
doi: 10.3389/fonc.2020.566637

Comparative Assessment of the WNT/β-Catenin Pathway, CacyBP/SIP, and the Immunoproteasome Subunit LMP7 in Various Histological Types of Renal Cell Carcinoma

Żaneta Piotrowska[1†], Michał Niezgoda[1†], Grzegorz Młynarczyk[2†], Magdalena Acewicz[1†] and Irena Kasacka[1†]*

[1] Department of Histology and Cytophysiology, Medical University of Białystok, Białystok, Poland, [2] Department of Urology, Medical University of Białystok, Białystok, Poland

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Ravi Manoharan,
University of Madras, India
Soichiro Yamamura,
University of California, San Francisco, United States

***Correspondence:**
Irena Kasacka
kasacka@umb.edu.pl;
histologia.cytofizjologia@umb.edu.pl

†ORCID:
Żaneta Piotrowska
orcid.org/0000-0003-2159-1068
Michał Niezgoda
orcid.org/0000-0003-4155-2621
Grzegorz Młynarczyk
orcid.org/0000-0002-1453-8501
Magdalena Acewicz
orcid.org/0000-0002-3398-366X
Irena Kasacka
orcid.org/0000-0003-2954-942X

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 28 May 2020
Accepted: 21 October 2020
Published: 20 November 2020

Citation:
Piotrowska Ż, Niezgoda M, Młynarczyk G, Acewicz M and Kasacka I (2020) Comparative Assessment of the WNT/β-Catenin Pathway, CacyBP/SIP, and the Immunoproteasome Subunit LMP7 in Various Histological Types of Renal Cell Carcinoma. Front. Oncol. 10:566637. doi: 10.3389/fonc.2020.566637

Objective: The Wnt/ß-catenin pathway plays an important role in pathogenesis of variety cancers. Most studies on changes in WNT/β-catenin pathway in renal cell carcinoma (RCC) apply only to clear cell RCC, while there are no comparative assessments of this signaling pathway in various histological types of renal tumors in the available literature. Additionally, considering the close relationship between WNT/β-catenin signaling, CacyBP/SIP and proteasomal activity, it seemed worth comparing WNT/β-catenin pathway, CacyBP/SIP and LMP7 immunoproteasome subunit in human samples of clear cell, papillary, and chromophobe RCC.

Methods: Tests were performed on sections of three types of kidney tumors together with surrounding unchanged tissue fragments collected from 50 patients. Samples were divided into three groups depending on the histological type of cancer: clear cell, papillary and chromophobe RCC. Immunohistochemistry and PCR methods were used to identify WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7, and gene expression.

Results: Immunoreactivity and expression of WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7 in clear cell RCC was markedly increased compared to non-cancerous kidney tissue. In papillary RCC, immunoreactivity and expression of WNT/β-catenin pathway, CacyBP/SIP, LMP7 was also increased compared to non-malignant kidneys, but it was less pronounced than in clear cell RCC. The least substantial increase in immunoreactivity and expression of WNT/β-catenin pathway, CacyBP/SIP, LMP7 was found in chromophobe RCC, compared to other RCC histological subtypes studied.

Conclusions: Study results suggest an important role of WNT/β-catenin pathway, CacyBP/SIP and LMP7 in RCC carcinogenesis, and may indicate new aspects of pathomechanisms leading to differences in the biology of clear cell, papillary and chromophobe RCC.

Keywords: WNT/β-catenin pathway, CacyBP/SIP, LMP7 immunosubunit, renal cell carcinoma, human

INTRODUCTION

Renal cell carcinoma (RCC) accounts for approximately 2% to 3% of all human cancers (1). The incidence of RCC is far higher in men who represent around two-thirds of all cases (1). Approximately 40% of RCC patients die of the disease, which makes it the third most deadly cancer among urological tumors (1). 25% to 30% of RCC patients are diagnosed with advanced metastatic disease and a further 30% to 50% of individuals with primary RCC develop metastases (2). The survival rate of patients with metastases is around 48% in the first year and only 9% five years after diagnosis (2). The most common histological subtypes of RCC are clear cell RCC (66–75% of cases), papillary RCC (10–15% of cases) and chromophobe RCC (5% cases) (1, 3). Clear cell RCC is associated with the highest invasiveness and shortest cancer-specific survival rate, while chromophobe RCC is related to the lowest aggressiveness and best prognosis compared to other RCC histological subtypes (1, 3, 4).

The Wnt/ß-catenin signaling pathway plays an important role in the process of carcinogenesis (5). It includes specific Frizzled membrane receptors (Fzd), endogenous agonists—WNT proteins as well as ß-catenin—the main effector of the pathway (5–7). ß-catenin is a bifunctional protein since it determines many aspects of cell physiology as a transcription factor, i.e. proliferation, differentiation, survival and migration, while as a membrane-bound protein, it coordinates intercellular adhesion and is responsible for maintaining tissue architecture (5–7). In the absence of binding of WNT ligands to Fzd receptors, intracellular ß-catenin levels are kept very low due to continuous degradation. The level of cytoplasmic ß-catenin is regulated by a multi-protein complex, called the ß-catenin destruction complex, consisting of axin, adenomatous polyposis coli (APC), casein kinase 1 (CK1) and glycogen synthase kinase 3 ß (GSK-3ß). In this complex, ß-catenin is phosphorylated by a dual-kinase mechanism involving CKI and GSK-3ß, and then phosphorylated ß-catenin is ubiquitinated and is directed to proteasomal degradation. Following the attachment of the WNT ligand and activation of Fzd receptors, destruction of ß-catenin is inhibited and the peptide accumulates in the cell. Then ß-catenin is transferred to the nucleus, where it regulates the expression of target WNT genes (5–7).

Increased activity of the WNT/β-catenin pathway has been found in various types of cancer such as colorectal cancer, hepatocellular carcinomas, melanoma, lung cancer, leukemia, bladder cancer, breast cancer, ovarian cancer and cervical cancer (8, 9). It has been demonstrated that the WNT/β-catenin pathway is involved in the pathogenesis and progression of RCC, tumor angiogenesis and invasive behavior of transformed cells (6). Several clinical and experimental studies have shown increased expression of WNT proteins and Fzd receptors, as well as increased content of β-catenin in RCC (6–12). Literature data also indicate shorter survival for RCC patients who have been found to have increased expression of WNT proteins or accumulated β-catenin in tumor tissue (9–11). Therefore, hyperactivation of WNT/β-catenin signaling may be of clinical significance in RCC (9–11). Genetic silencing of the WNT/β-catenin pathway in RCC cell lines has been demonstrated to result in inhibition of proliferation and migration, induction of apoptosis and increase of cancer cells sensitivity to chemotherapy (11).

Some recent evidence suggests that GSK-3ß, a key kinase of the aforementioned β-catenin destruction complex, may also be a promising therapeutic target in RCC (13). Increased GSK-3ß content has been found in human RCC samples and cultured RCC cell lines (13–15). Genetic or pharmacological inhibition of GSK-3ß contributes to cell cycle arrest and reduces viability of RCC cell lines (13–15).

An alternative pathway of β-catenin degradation with participation of the CacyBP/SIP protein (Calcyclin-binding protein/Siah-1–interacting protein) has recently been discovered (16). CacyBP/SIP is a multi-domain protein that interacts with a wide range of intracellular molecules, including ubiquitin ligases components Siah-1 (seven in absentia homolog-1) and Skp1 (S-phase kinase associated protein 1). CacyBP/SIP attaches to Siah-1 and Skp1, stabilizes the ubiquitin ligase complex and promotes the degradation of non-phosphorylated β-catenin (16). Independent scientific centers have indicated the importance of CacyBP/SIP in the process of cancer formation. Abnormal CacyBP/SIP levels have been found in pancreatic cancer, gastric cancer, colorectal cancer, osteogenic sarcoma, melanoma, kidney cancer, breast cancer, brain cancer (16).

It has been shown that CacyBP/SIP can act as a tumor suppressor or an oncogen depending on the type of cancer (16–18). Decreased CacyBP/SIP expression has been found in samples taken from patients with RCC as well as in several RCC lines (19). Forced overexpression of CacyBP/SIP in RCC cell lines has been shown to reduce the proliferative potential of cancer cells *in vitro* and their carcinogenicity when injected into mice (19).

The body's immune system constantly attacks and destroys cancer cells that form. Disturbances in antigen presentation on cell surfaces have been found in various types of cancers, which may be a mechanism for avoiding recognition and destruction by the immune system (20–22). The crucial role in antigen processing is played by immunoproteasomes involved in the production of antigenic peptides and the major histocompatibility complex class I (MHC I), which is responsible for presentation of antigens to immunologically competent cells (20–22). Studies on patients with RCC have shown reduced expression of both immunoproteasomes and MHC I in malignant kidney samples. These findings indicate that in RCC cells, tumor antigen presentation may be inhibited in order to avoid an anti-tumor immune response (20–22).

Most studies on changes in the WNT/β-catenin pathway in RCC focus only on clear cell RCC, while there is no published literature on the comparative assessment of this signaling pathway in different histological types of RCC (6–12). Additionally, considering the close relationship between WNT/β-catenin signaling, CacyBP/SIP and proteasomal activity, it seemed worth comparing the WNT/β-catenin pathway, CacyBP/SIP and LMP7 immunoproteasome subunit in human samples of clear cell, papillary and chromophobe RCC.

The aim of this study is immunohistochemical evaluation and comparison of the expression of the WNT10A, Fzd5, β-catenin,

GSK-3ß, CacyBP/SIP, and LMP7 genes in human kidney cancer tissue in different histological types of RCC.

MATERIAL AND METHODS

Sample Collection

The research was conducted on postoperative material collected from fifty patients of Department of Urology, Medical University of Bialystok, operated on for kidney cancer. The study protocol was approved by the Bioethics Committee, Medical University of Bialystok (R-I-002/282/2019) and prior written informed consent was obtained from each subject.

The research material were fragments of RCC lesions obtained during radical or partial nephrectomy. The comparative material were fragments of surrounding unchanged kidney tissue (margins). All RCC lesions were at the same grade G2, except for chromophobe RCC in which grading system is not applicable. Obtained malignant and non-cancerous kidney tissues were immediately fixed in Bouin's solution and routinely embedded in paraffin or placed in RNA-later solution (AM7024 Thermo Fischer) and stored in −80°C.

The subjects were divided into 3 groups, depending on the histological type of cancer: clear cell RCC (35 cases), papillary RCC (12 cases) and chromophobe RCC (3 cases).

Renal paraffin blocks were cut into 4 μm section thickness and then stained with hematoxylin-eosin for general histological examination and processed by immunohistochemistry to detect WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7. Material stored in RNA-later solution was processed by real-time PCR to evaluate the expression of genes coding WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7.

Immunohistochemistry

In the immunohistochemical study, the EnVision method was used according to Herman and Elfont (23). Immunohistochemistry was performed, using an REAL™ EnVision™ Detection System, Peroxidase/DAB, Rabbit/Mouse detection kit (K5007; DakoCytomation; Glostrup, Denmark). Immunostaining was performed by the following protocol: paraffin-embedded sections were deparaffinized and hydrated in pure alcohols. For antigen retrieval, the sections were subjected to pretreatment in a pressure chamber heated for 1 min at 21 psi (one pound force per square inch (1 psi) equates to 6.895 kPa, the conversion factor has been provided by the United Kingdom National Physical Laboratory) at 125°C, using Target Retrieval Solution Citrate pH=6.0 S 2369 (DakoCytomation; Glostrup, Denmark) for WNT10A, Fzd5, β-catenin, CacyBP/SIP and LMP7 and Target Retrieval Solution pH = 9.0 for GSK-3ß. After cooling down to room temperature, the sections were incubated with Peroxidase Blocking Reagent S 2001 (DakoCytomation; Glostrup, Denmark) for 10 min to block endogenous peroxidase activity. Subsequently sections were incubated with primary antibody for WNT10A (Rabbit polyclonal to Wnt10a ab106522 Abcam), Fzd5 (Rabbit polyclonal to Frizzled 5 ab115204 Abcam), β-catenin (Recombinant Anti-beta Catenin antibody (E247) ab32572 purchased from Abcam), GSK-3ß (Rabbit monoclonal (EPR933Y) to GSK3 (alpha + beta) (phospho Y216 + Y279) ab68476 Abcam), CacyBP/SIP (Rabbit polyclonal to CacyBP ab190950 Abcam) and LMP7 (mouse monoclonal antibody to LMP7, PW8845, purchased from Biomol). All antibodies were previously diluted in Antibody Diluent Background Reducing (S 3022 DakoCytomation; Glostrup, Denmark) in relation 1:1000 for WNT10A antibody, 1:150 for Fzd5 antibody, 1:2000 for β-catenin antibody, 1:100 for GSK-3ß antibody, 1:600 for CacyBP/SIP antibody and 1:5000 for LMP7 antibody. Sections with WNT10A-, Fzd5-, β-catenin-, GSK-3ß-, CacyBP/SIP-, and LMP7-antibody were incubated overnight at 4°C (incubation performed in a humidified chamber). The procedure was followed by incubation with secondary antibody (conjugated to horseradish peroxidase-labeled polymer). The bound antibodies were visualized by 1-min incubation with liquid 3,3′-diaminobenzidine substrate chromogen. The sections were finally counterstained in hematoxylin QS (H-3404, Vector Laboratories; Burlingame, CA), mounted, covered and evaluated under a light microscope. Appropriate washing with Wash Buffer (S 3006 DakoCytomation; Glostrup, Denmark) was performed between each step.

Specificity tests performed for WNT10A-, Fzd5-, β-catenin-, GSK-3ß-, CacyBP/SIP, and LMP7 included a negative control in which the antibodies were replaced by normal rabbit serum (Vector Laboratories; Burlingame, CA) with appropriate dilution. All these controls were negative.

Histological preparations were visually analyzed using an Olympus BX43 light microscope (Olympus 114 Corp., Tokyo, Japan) with an Olympus DP12 digital camera (Olympus 114 Corp., Tokyo, Japan) and documented.

Quantitative Analysis

Twelve sections of malignant lesion and twelve section of non-cancerous kidney tissue were examined from each subject (two section for each WNT10A-, Fzd5-, β-catenin-, GSK-3ß-, CacyBP/SIP- and LMP7-immunostaining). Five randomly selected microscopic fields (each field 0.785 mm^2, 200× magnification (20× lens and 10× eyepiece)) from each kidney section were documented using an Olympus DP12 microscope camera. Each obtained digital image of the kidney section was morphometric evaluated using NIS Elements AR 3.10 Nikon software for microscopic image analysis.

The intensity of the immunohistochemical reaction for all the antibodies used in the study was measured on each image analyzed and determined using a gray scale level 0 to 255, where the value of the completely white or bright pixel is 0, while the completely black pixel is 255.

Real-Time PCR

Samples of kidney cancer and non-malignant renal tissue (1 cm^3) were taken from each patient and placed in an RNA-later solution. Total RNA was isolated using NucleoSpin® RNA Isolation Kit (Machery-Nagel). Quantification and quality control of total RNA was determined using a spectrophotometer - NanoDrop 2000 (ThermoScientific). An aliquot of 1 μg of total RNA was reverse transcribed into cDNA using iScript™ Advanced cDNA Synthesis

Kit for RT-qPCR (BIO-RAD). Synthesis of cDNA was performed in a final volume of 20 µl using an Thermal Cycler (Model SureCycler 8800, Aligent Technologies). For reverse transcription, the mixtures were incubated at 46°C for 20 min, then heated to 95°C for 1 min and finally cooled quickly at 4°C. Quantitative real-time PCR reactions were performed using Stratagene Mx3005P (Aligent Technologies) with the SsoAdvanced™ Universal SYBER® Green Supermix (BIO-RAD). Specific primers for WNT10A (*WNT10A*), Fzd5 (*FZD5*), β-catenin (*CTNNB1*), GSK-3ß (*GSK3B*), CacyBP/SIP (*CACY BP*), LMP7 (*PSMB8*), and GAPDH (*GAPDH*) were designed by BIO-RAD Company. The housekeeping gene GAPDH (*GAPDH*) was used as a reference gene for quantification. To determine the amounts of levels of test genes expression, standard curves were constructed for each gene separately with serially diluted PCR products. PCR products were obtained by cDNA amplification using specific primers as follows: *WNT10A* (qHsaCED0044646, BIO-RAD), *FZD5* (qHsaCJD0034589, BIO-RAD), *CTNNB1* (qHsaCED0046518, BIO-RAD), *GSK3B* (qHsaCED0057061, BIO-RAD), *CACY BP* (qHsaCED0043669, BIO-RAD), *PSMB8* (qHsaCED0037294, BIO-RAD), and *GAPDH* (qHsaCED0038674, BIO-RAD). qRT-PCR was carried out in a doublet in a final volume of 20 µl under the following conditions: 2 min polymerase activation at 95°C, 5 s denaturation at 95°C, 30 s annealing at 60°C for 35 cycles. PCR reactions were checked, including no-RT-controls, omitting of templates, and melting curve to ensure only one product was amplified. The relative quantification of gene expression was determined by comparing Ct values using the ΔΔCt method. All results were normalized to GAPDH.

Statistical Analysis

All data were analyzed for statistical significance using the Statistica version 12.0 computer software package. The mean values were computed automatically; significant differences were determined by one-way ANOVA test; $p < 0.05$ was considered significant.

RESULTS

Baseline characteristics of patients with RCC, including the proportion of men and women, mean values of age, minimal and maximal value of age are presented in **Table 1**. Men consisted almost two-thirds of studied RCC cases. Patients with RCC belonged to middle-age group, most of the cases concerned individuals in the fifth or sixth decade of life (**Table 1**).

TABLE 1 | Characteristic of patients with clear cell, papillary and chromophobe RCC.

Group of patients	Proportion of men	Proportion of women	Age [years]		
			Mean	Min	Max
Overall	32/50	18/50	66.2 ± 1.30	40	87
Clear cell RCC	20/35	15/35	67.0 ± 1.32	43	87
Papillary RCC	10/12	2/12	62.1 ± 6.10	40	82
Chromophobe RCC	2/3	1/3	59.3 ± 2.03	56	63

Immunohistochemistry

A positive immunohistochemical reaction for WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7 was observed in all studied kidneys tissues, although the intensity of immunoreaction varied between non-malignant kidney and RCC tissue samples, depending on the histological type of the tumor (**Figures 1–6**).

In an unchanged kidney, there was a weak or moderate immune response against WNT10A in the renal tubules and very delicate WNT10A immunostaining in the glomeruli (**Figure 1A**). Considerably stronger immunoreactivity for WNT10A was found in clear cell RCC compared to non-malignant tissue (**Figure 1B**). In papillary RCC, WNT10A immunoreactivity was also increased compared to non-cancerous kidney, but to a lesser degree than in clear cell RCC (**Figure 1C**). There were no noticeable differences in the intensity of WNT10A immunostaining between chromophobe RCC and non-malignant tissue (**Figure 1D**).

Immunodetection of Fzd5 in non-malignant kidney resulted in medium or bright brown staining of the renal tubules and weak reactivity in the glomeruli (**Figure 2A**). The intensity of reaction against Fzd5 was significantly higher in RCC tissue (**Figures 2B–D**) compared to unchanged kidney. The highest immunoreactivity for Fzd5 was observed in clear cell RCC (**Figure 2B**), while the lowest increase in Fzd5 intensity was observed in chromophobe RCC in comparison to other histological types of RCC studied (**Figure 2D**).

Immunohistochemistry showed the presence of β-catenin in the renal tubules of non-cancerous kidney tissue. Distal tubules showed moderate β-catenin immunolabeling, while in proximal tubules the reaction against β-catenin was weak (**Figure 3A**). In clear cell RCC, significantly increased immunoreactivity of β-catenin was observed compared to non-cancerous control. The β-catenin immunosignal was very strong and located mainly in the peripheral area of malignant cells cytoplasm (**Figure 3B**). A significant increase in β-catenin immunoreaction was also observed in papillary RCC (**Figure 3C**), although the observed intensification of β-catenin reactivity was lower than in clear cell RCC. In chromophobe RCC, the intensity of β-catenin staining was also stronger compared to unchanged renal tissue, but the observed change in β-catenin immunoreactivity was least pronounced in comparison with other studied histological types of RCC (**Figure 4D**).

Antisera against GSK-3ß moderately immunostained renal tubules and glomeruli in non-malignant kidney (**Figure 4A**). GSK-3ß immunoreaction was noticeably increased in clear cell RCC (**Figure 4B**) compared to non-malignant kidney. In papillary RCC (**Figure 4C**) and chromophobe RCC (**Figure 4D**), intensity of GSK-3ß immunoreactivity was lower compared to clear cell RCC and did not differ from that observed in unchanged kidney.

CacyBP/SIP immunostaining in non-cancerous kidney revealed its location in the renal tubule system and glomeruli. In the distal tubules, moderate immunoreaction against CacyBP/SIP was observed, whereas in the proximal tubules and glomeruli, immunoreactivity for CacyBP/SIP was weak (**Figure 5A**). A much

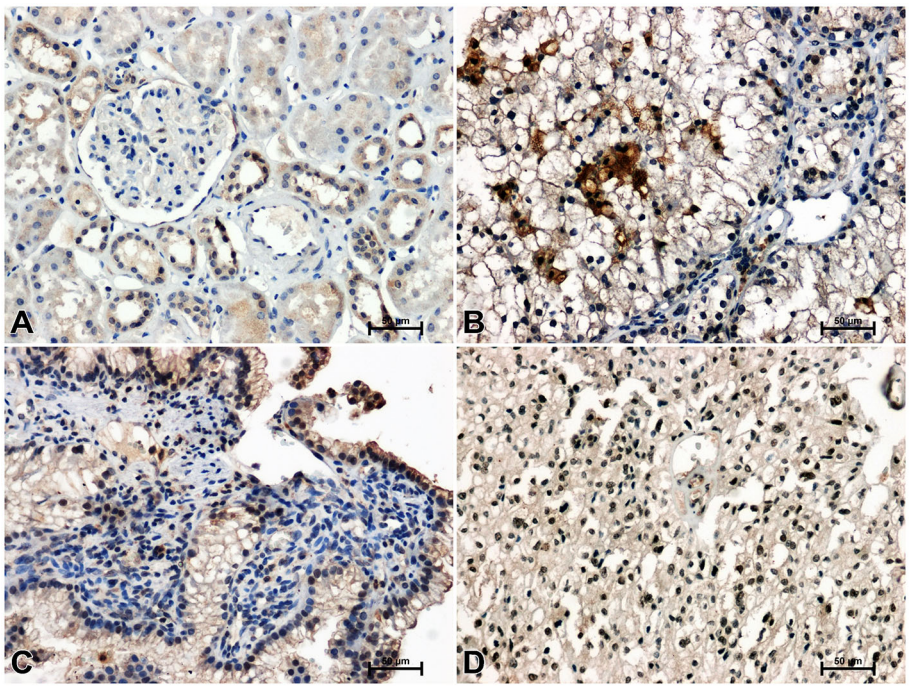


FIGURE 1 | Immunoidentification of WNT10A in non-malignant kidney **(A)** and clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

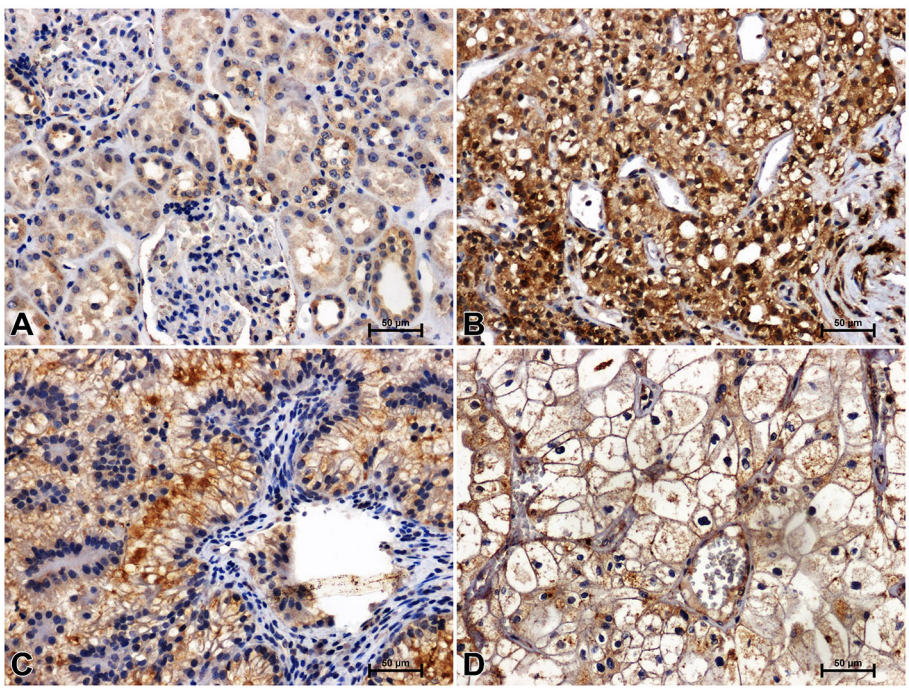


FIGURE 2 | Immunodetection of Fzd5 in non-malignant kidney **(A)** and clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

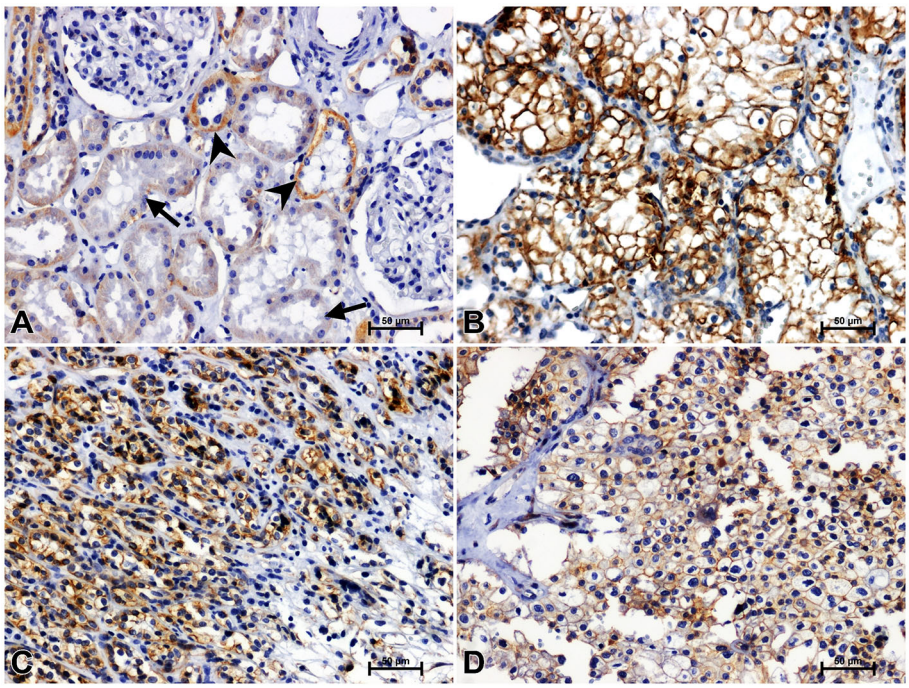


FIGURE 3 | Positive β-catenin immunostaining in distal (arrowheads) and proximal (arrows) tubules of non-malignant kidney **(A)**. Results of β-catenin immunostainig in clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

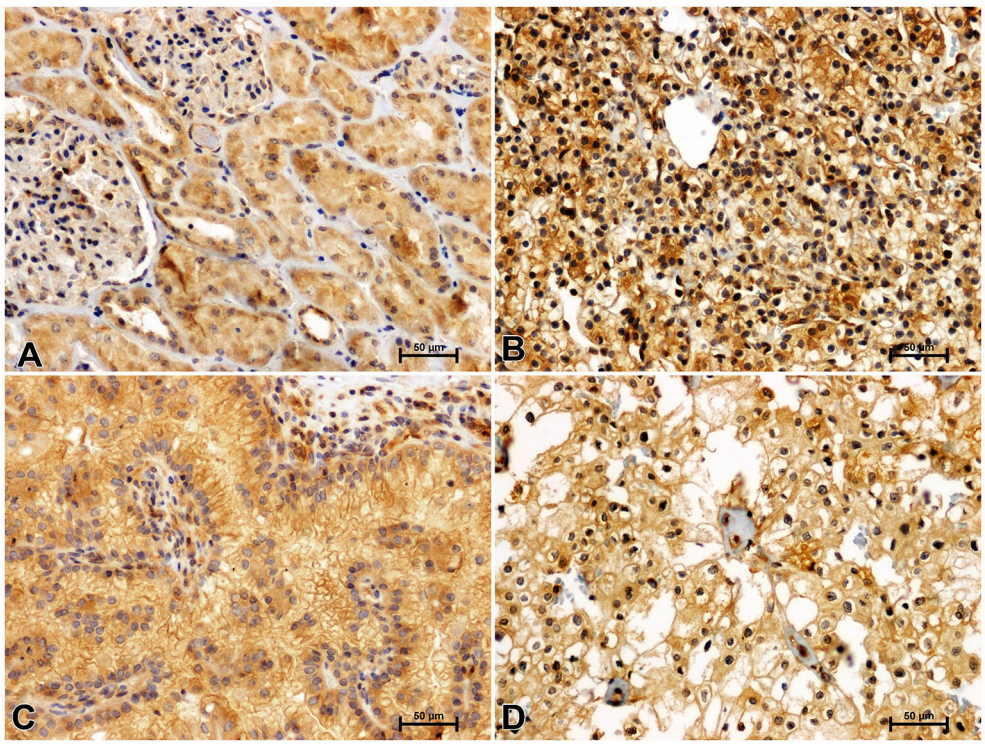


FIGURE 4 | Immunolabeling of GSK-3ß in non-malignant kidney **(A)** and clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

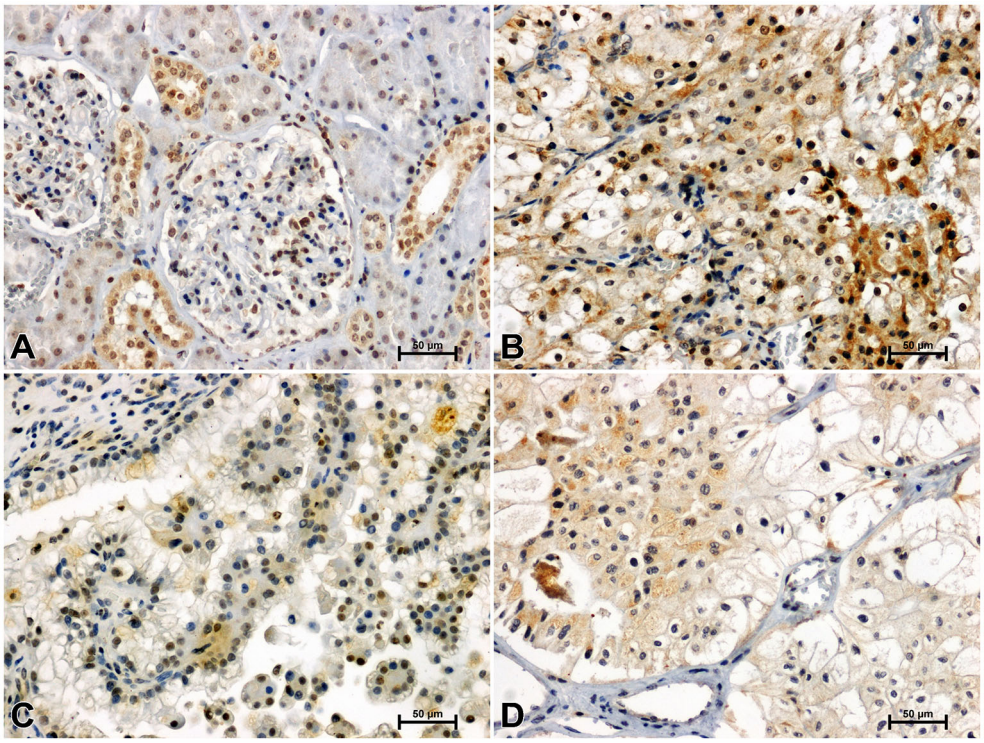


FIGURE 5 | Immunohistochemical reaction determining CacyBP/SIP in non-malignant kidney **(A)** and clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

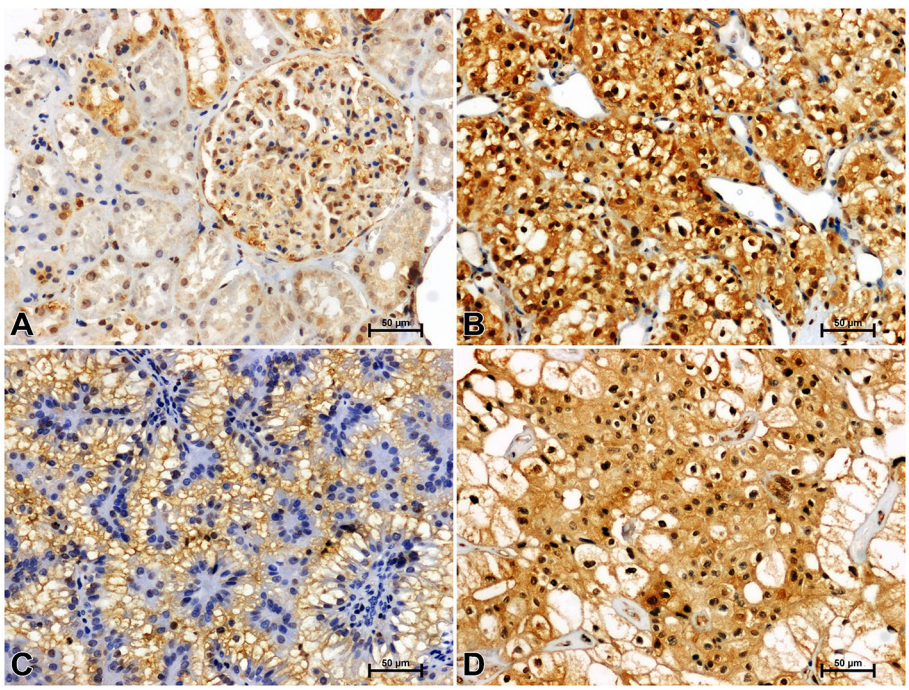


FIGURE 6 | Representative immunohistochemical photomicrographs of LMP7 in non-malignant kidney **(A)** and clear cell RCC **(B)**, papillary RCC **(C)**, chromophobe RCC **(D)**.

stronger intensity of the CacyBP/SIP immunosignal was observed in clear cell RCC, as compared to unchanged renal tissue (**Figure 5B**). Immunohistochemical reaction against CacyBP/SIP in papillary RCC (**Figure 5C**) and chromophobe RCC (**Figure 5D**) was weaker compared to clear cell RCC and was similar to that in non-cancerous control tissue.

Immunoidentification of LMP7 in non-cancerous kidney tissue gave a weak response in the proximal tubules and moderate reactivity in the distal tubules and glomeruli (**Figure 6A**). Greater immunoreactivity for LMP7 was noted in clear cell RCC compared to unchanged kidney (**Figure 6B**). Immunostaining of LMP7 in papillary RCC (**Figure 6C**) and chromophobe RCC (**Figure 6D**) was increased compared to non-malignant renal tissue, although it was lower than in clear cell RCC.

Quantitative Analysis

The results of densimetric tests confirmed visually perceptible differences in the intensity of immunoreaction against WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7 between studied histological types of RCC and non-cancerous control samples (**Table 2**).

Real-Time PCR

QRT-PCR analysis revealed a significant increase in the expression of the WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7 genes in clear cell RCC compared to non-malignant kidney. Similarly, noticeably increased expression of the tested genes was found in papillary RCC compared to unchanged renal tissue, but the observed increase was smaller than in clear cell RCC. The weakest expression of the analyzed genes was observed in chromophobe RCC compared to other studied histological types of RCC. Expression of WNT10A, GSK-3ß, and LMP7 in chromophobe RCC was marginally elevated compared to non-cancerous control, while expression of the Fzd5, β-catenin and CacyBP/SIP genes remained at the level observed in unchanged kidney (**Figure 7** and **Table 3**).

DISCUSSION

According to the latest data from the International Agency for Research on Cancer, approximately 403,000 new cases of kidney cancer were diagnosed globally in 2018 (24). Kidney cancer is the ninth and fourteenth most common malignancy in men and women, respectively (25). Each year, between 150 and 175,000 patients die from this cancer, which ranks it third in terms of mortality among urological cancer (1, 24, 25). Advancing our knowledge of the mechanisms involved in RCC pathogenesis is necessary for improving the effectiveness of prevention and treatment of this malignant disorder.

Considering the important role of the Wnt/ß-catenin pathway in carcinogenesis and the close relationship between WNT/β-catenin signaling, CacyBP/SIP and proteasomal activity, it seemed worthwhile to undertake research aimed at comparing the WNT/β-catenin pathway, CacyBP/SIP and LMP7 immunoproteasome in human samples of different RCC histological types.

In presented study were examined 50 patients with various histological types of RCC: 35 with clear cell RCC, 12 with papillary RCC, 3 with chromophobe RCC. Men represented almost two-third of the studied RCC cases, which is consistent with literature data indicating on higher incidence of RCC in men than in women (1). Our study showed markedly increased immunoreactivity for WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, and LMP7 in clear cell RCC compared to unchanged kidney tissue. In papillary RCC, immunoreactivity of the tested antibodies (except GSK-3ß and CacyBP/SIP) was also increased compared to non-malignant kidney, but it was weaker than in clear cell RCC. Reaction intensity for GSK-3ß and CacyBP/SIP in papillary RCC was comparable to that observed in non-malignant kidney. Immunoreactivity of the tested antibodies in chromophobe RCC was at the level observed in non-cancerous kidney cells and only for the Fzd5, β-catenin and LMP7 antibodies was it marginally stronger compared to control samples. Gene expression analysis confirmed the results obtained in immunohistochemical reactions.

Alterations in Wnt/ß-catenin pathway activity in RCC have been revealed in several previous investigations. Hsu et al. (11) demonstrated significantly increased expression of the WNT10A gene in RCC lines and in tissues of human RCC compared to non-cancerous control samples. In subsequent immunohistochemical studies, the authors also found increased content of WN10A and β-catenin in human RCC compared to normal renal tissue (11).

TABLE 2 | The intensity of immunoreaction determining WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7 in clear cell RCC, papillary RCC, chromophobe RCC, and non-malignant renal tissue (mean ± SE).

Group of patients	Intensity of immunohistochemical reaction in kidneyScale from 0 (white pixel) to 255 (black pixel)					
	WNT10A	Fzd5	β-catenin	GSK-3ß	CacyBP/SIP	LMP7
Control	76.0 ± 2.93	95.1 ± 2.01	101.4 ± 2.17	107.4 ± 2.17	86.2 ± 3.40	95.5 ± 154.9
Clear cell RCC	152.2 ± 4.62$^{a\uparrow}$	185.0 ± 2.60$^{a\uparrow}$	167.1 ± 2.71$^{a\uparrow}$	144.6 ± 2.90$^{a\uparrow}$	131.1 ± 3.58$^{a\uparrow}$	154.9 ± 3.04$^{a\uparrow}$
Papillary RCC	113.5 ± 5.00$^{a\uparrow\ b\downarrow}$	154.9 ± 5.42$^{a\uparrow\ b\downarrow}$	134.6 ± 3.39$^{a\uparrow\ b\downarrow}$	110.5 ± 2.35$^{b\downarrow}$	79.7 ± 2.10$^{b\downarrow}$	114.4 ± 2.19$^{a\uparrow\ b\downarrow}$
Chromophobe RCC	73.7 ± 2.30$^{c\downarrow\ d\downarrow}$	121.1 ± 2.82$^{a\uparrow\ c\downarrow\ d\downarrow}$	119.8 ± 3.03$^{a\uparrow\ c\downarrow\ d\downarrow}$	107.1 ± 2.55$^{c\downarrow}$	79.7 ± 2.63$^{c\downarrow}$	114.3 ± 2.46$^{a\uparrow\ c\downarrow}$

$^{a}p < 0.05$—RCC vs control kidney.
$^{b}p < 0.05$—papillary RCC vs clear cell RCC.
$^{c}p < 0.05$—chromophobe RCC vs clear RCC.
$^{d}p < 0.05$—chromophobe RCC vs papillary RCC.
↑—intensification of immunohistochemical reaction.
↓—weakening of immunohistochemical reaction.

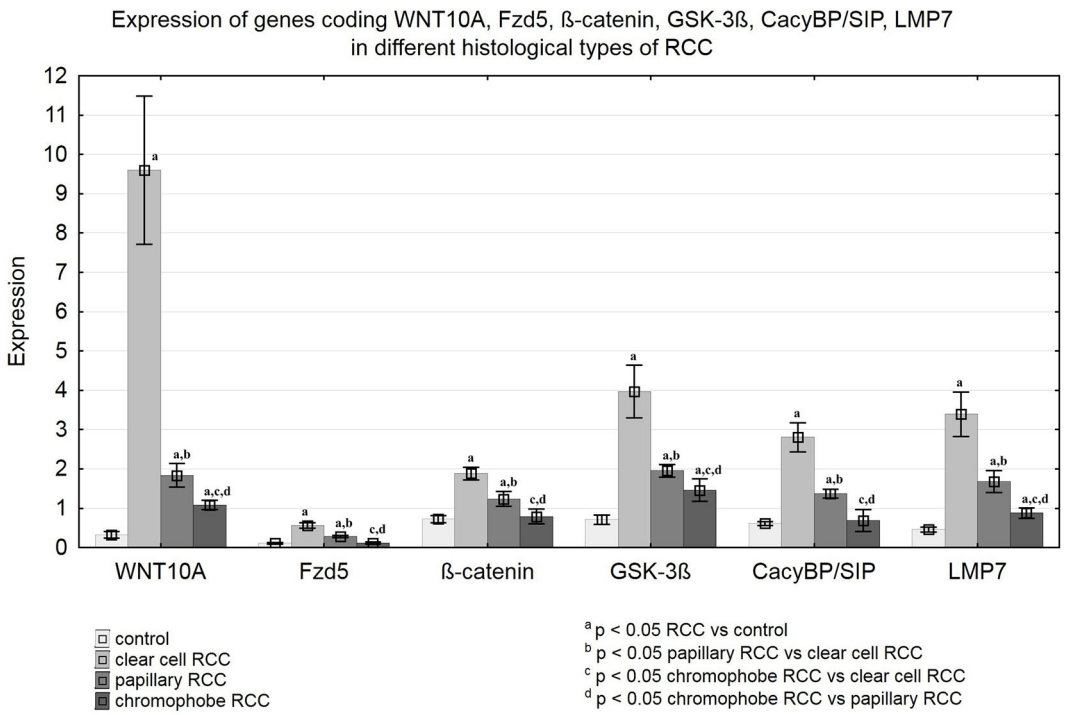


FIGURE 7 | Expression of genes coding WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7 in clear cell RCC, papillary RCC, chromophobe RCC and non-malignant renal tissue.

TABLE 3 | The expression of genes coding WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7 in clear cell RCC, papillary RCC, chromophobe RCC and non-malignant renal tissue (mean ± SE).

Group of Patients	The expression of genes coding WNT10A, Fzd5, β-catenin, GSK-3ß, CacyBP/SIP, LMP7 in kidney					
	WNT10A	**Fzd5**	**β-catenin**	**GSK-3ß**	**CacyBP/SIP**	**LMP7**
Control	0.32 ± 0.084	0.11 ± 0.019	0.72 ± 0.086	0.71 ± 0.117	0.61 ± 0.004	0.46 ± 0.060
Clear cell RCC	9.60 ± 1.889$^{a\uparrow}$	0.56 ± 0.069$^{a\uparrow}$	1.88 ± 0.163$^{a\uparrow}$	3.97 ± 0.666$^{a\uparrow}$	2.81 ± 0.374$^{a\uparrow}$	3.39 ± 0.562$^{a\uparrow}$
Papillary RCC	1.84 ± 0.300$^{a\uparrow\ b\downarrow}$	0.28 ± 0.033$^{a\uparrow\ b\downarrow}$	1.24 ± 0.189$^{a\uparrow\ b\downarrow}$	1.95 ± 0.164$^{a\uparrow\ b\downarrow}$	1.37 ± 0.111$^{a\uparrow\ b\downarrow}$	1.68 ± 0.276$^{a\uparrow\ b\downarrow}$
Chromophobe RCC	1.08 ± 0.115$^{a\uparrow\ c\downarrow\ d\downarrow}$	0.11 ± 0.001$^{c\downarrow\ d\downarrow}$	0.79 ± 0.185$^{c\downarrow\ d\downarrow}$	1.46 ± 0.280$^{a\uparrow\ c\downarrow\ d\downarrow}$	0.69 ± 0.275$^{c\downarrow\ d\downarrow}$	0.88 ± 0.130$^{a\uparrow\ c\downarrow\ d\downarrow}$

$^{a}p < 0.05$—RCC vs control kidney.
$^{b}p < 0.05$—papillary RCC vs clear cell RCC.
$^{c}p < 0.05$—chromophobe RCC vs clear RCC.
$^{d}p < 0.05$—chromophobe RCC vs papillary RCC.
↑—increased expression of gene.
↓—decreased expression of gene.

Janssens et al. (12) revealed increased expression of some Fzd receptors in human RCC samples, including Fzd5 compared to a healthy kidney. Subsequent Western blot analysis confirmed the increase in Fzd5 content in kidney cancer compared to a healthy kidney (12). Liu et al. (9) demonstrated enhanced levels of β-catenin in the blood of patients with RCC in comparison to healthy individuals. Similarly, the authors noted increased expression of the β-catenin gene and elevated peptide content in RCC samples compared to non-malignant renal tissue (9). Other studies have shown elevated level of GSK-3ß in human RCC samples and cultured RCC cell lines, as compared to non-cancerous control samples (13). The reports mentioned above describe changes in WNT/β-catenin signaling only in clear cell RCC. To the best of our knowledge, the present study is the first comparative assessment of this signaling pathway in various histological types of RCC.

The proliferative potential of cancer cells reflects tumor aggressiveness and is strongly correlated with clinical prognosis. Mehdi et al. (26) examined proliferation rates in various histological types of RCC using Ki-67 and mini-chromosome maintenance 2 (MCM-2) markers. The study showed an increased level of proliferative markers in all samples of RCC, wherein in clear cell RCC the proliferation index was the most enhanced compared to papillary RCC and

chromophobe RCC (26). Several *in vitro* studies have demonstrated that the WNT/β-catenin pathway plays an important role in regulating the proliferation of kidney cancer cells. Hsu et al. (11) stated that forced overexpression of WNT10A in RCC cell lines resulted in enhanced proliferation and invasion of malignant cells, while silencing WNT10A and the β-catenin gene leads to cell cycle arrest and reduced renal cancer cell migration. Yang et al. (27) genetically suppressed the expression of Fzd8 receptors in RCC cell lines and noticed lower proliferative potential in cancer cells with knocked-down Fzd8 gene compared to those without modification. Von Schulz-Hausmann et al. (28), Koller, et al. (29), and Sasamura et al. (30) demonstrated that incubation of renal cancer cell cultures with WNT/β-catenin signaling inhibitors slowed the proliferation of neoplastic cells. Bilim et al. (13), Kawazoe et al. (14) and Pal et al. (15) found that inactivation of the GSK-3ß gene or treatment with GSK-3ß inhibitors reduced the proliferation of cultured RCC cells. In subsequent studies, the scientists injected renal cancer cells into mice subcutaneously and investigated the effect of pharmacological inhibition of GSK-3ß on tumor growth. They noticed that in mice that received GSK-3ß inhibitors, the tumor grew slower than in untreated animals (13–15). Considering the above, greater proliferative potential of cancer cells in clear cell RCC compared to other RCC histological subtypes might be associated with the highest increase in the WNT/β-catenin pathway observed in clear cell RCC as compared to papillary RCC and chromophobe RCC.

WNT/β-catenin signaling is also implicated in controlling renal cancer cell survival. Activation of the WNT/β-catenin pathway inhibits cell apoptosis by reducing the release of cytochrome c from mitochondria and decreasing caspase levels (31). RCC cells with forced WNT10A overexpression have been shown to be more resistant to the cytotoxic effects of chemotherapy drugs (11). Conversely, suppression of WNT10A and β-catenin expression in RCC cells increases their susceptibility to chemotherapy (11). Silencing the Fzd8 receptor gene in cultured RCC cells significantly reduces their viability compared to cells with unmodified Fzd8 receptor expression (27). Similarly, treatment of RCC cells with WNT/β-catenin signaling inhibitors negatively affects their vitality (28–30). Genetic or pharmacological inhibition of GSK-3ß in RCC cells leads to cancer cell apoptosis, causing disturbances in intracellular energy homeostasis, downregulation of anti-apoptotic proteins and activation of caspases (13–15). Our research results indicate that in RCC, the WNT/β-catenin signaling pathway is activated with varying degrees of intensity depending on the type of cancer. Our findings may suggest differences in the regulation of signaling pathways determining the survival of malignant cells in examined histological subtypes of RCC.

Clinical data indicate far more frequent occurrence of metastases in clear cell RCC and papillary RCC in comparison to chromophobe RCC (3, 4). Neoplastic cells become invasive and spread from the primary site when they lose their intercellular connection and detach from surrounding cells (32). The basic components of cell-cell junctions are cadherins (32). Langner et al. (33), Kuehn et al. (34) and Shen et al. (35) evaluated immunohistochemically cadherin expression in human RCC tissues. These studies revealed positive immunostaining for cadherins only in a small percentage of clear cell RCC and papillary RCC samples, while in almost all cases of chromophobe RCC immunoreactivity for cadherins was observed (33–35). Gerharz et al. (36) described the ultrastructural aspects of cancer cells in various histological types of RCC and found that desmosomal junctions were rarely seen in clear cell RCC and papillary RCC, while in chromophobe RCC, typical desmosomes density was observed. The above findings indicate relatively undisturbed interactions between cells in chromophobe RCC, and significant disruption in the intercellular junction in clear cell RCC and papillary RCC. β-catenin plays a fundamental role in two different cellular processes: Wnt-mediated transcription activation and coordination of intercellular junctions by forming complexes with membrane cadherins (37). Proper balancing between the transcriptional and adhesive activity of β-catenin is crucial for maintaining cell function (37). In the state of chronic overactivation of the WNT/β-catenin pathway, β-catenin may be directed to transcriptional gene regulation at the expense of binding β-catenin to cadherin and its participation in intercellular adhesion. In view of the above, it might be suspected that, there are less severe disorders in intercellular junctions in chromophobe RCC compared to clear cell RCC and papillary RCC due to different intensity of changes in WNT/β-catenin signaling in the studied histological subtypes of RCC. Hence, different degrees of alteration in WNT/β-catenin signaling between clear cell RCC, papillary RCC and chromophobe RCC revealed in our study can provide a possible explanation for the distinct invasive behavior of the examined histological types of RCC.

The histological RCC subtype is an important clinical prognostic factor. In patients with clear cell RCC, prognosis is worse than in patients with papillary RCC and chromophobe RCC. The 5-year survival rate in clear cell RCC is 43% to 89%, while in papillary RCC and chromophobe RCC, it is 57% to 86% and 76% to 100%, respectively (3, 4). Some recent evidence has demonstrated a significant relationship between WNT/β-catenin signaling and the survival of patients with RCC. It has been revealed that higher serum β-catenin levels and increased content of WNT10 and β-catenin in cancer tissue are associated with higher mortality due to RCC (9, 11). It may be hypothesized that greater activation of the WNT/β-catenin pathway in clear cell RCC in comparison to papillary RCC and chromophobe RCC, as stated in the present report, might be the mechanism responsible for a worse prognosis in clear cell RCC compared to other examined histological subtypes of RCC.

Recent literature data indicates abnormal levels of CacyBP/SIP in various types of cancer such as gastric cancer, breast cancer, pancreatic cancer, colorectal cancer and brain cancer (16–18). However, there is only one report in the available literature on changes in CacyBP/SIP in renal cancer and therefore the role of this peptide in RCC cancerogenesis is not fully understood (19). Sun et al. (19) showed decreased CacyBP/SIP immunoreactivity and reduced CacyBP/SIP gene expression in human RCC samples

and RCC cell lines compared to non-cancerous control tissue samples. On the contrary, our study revealed an intensified CacyBP/SIP immunostaining and increased CacyBP/SIP gene expression in human RCC tissue compared to non-cancerous kidney. This discrepancy may be due to differences in the methods used. The role of CacyBP/SIP in cancer progression appears to be contradictory – it has been shown to either promote or inhibit the proliferation and invasiveness of cancer cells, depending on the type of cancer (16–19). Investigation by Sun et al. (19) demonstrated, that injecting mice with CacyBP/SIP overexpressing RCC cells reduces their proliferative potential and carcinogenicity. Similar tumor-suppressory effect of CacyBP/SIP was stated in studies on gastric cancer, breast cancer and astrocytoma. On the other hand oncogene role of CacyBP/SIP was demonstrated in pancreas cancer and glioma. In colorectal cancer, CacyBP/SIP has been shown to reduce proliferation while enhancing the invasion and migration of cancer cells (16, 18). Considering current incomplete state of knowledge about role of CacyBP/SIP in RCC cancerogenesis, more research needs to be done to determine precisely the consequences of CacyBP/SIP alterations in RCC cancer.

Recent investigations have shown that in RCC cancer interferes the presentation of antigens on the cell surface to avoid being recognized and destroyed by the immune system (20–22). The antigen presentation process involves immunoproteasomes that produce antigenic peptides and MHC I complexes that exhibit surface antigens for immunologically competent cells (20–22). Seliger et al. (20, 21) and Atkins et al. (22) demonstrated significant downregulation of MHC complexes in human RCC. The authors also demonstrated weakened or negative LMP7 immunosignaling in a large proportion of human RCC samples studied, revealing immunoproteasomes expression deficiency in RCC cancer (20–22). It has been proposed that observed abnormalities in the expression of MHC I and LMP7 in RCC tissue contribute to the suppression of tumor antigen presentation and may be a mechanism of immune escape in RCC cancer (20–22).

In contrast to the findings presented above, our report showed stronger immunostaining for LMP7 and increased expression of the LMP7 gene in RCC human tissue compared to a healthy kidney. Malignant cells have uncontrolled proliferative potential and enhanced metabolism, and therefore are more likely to accumulate defective proteins (38). Therefore, accelerated proteasomal activity has been found in many types of cancer in response to increased demand for protein turnover (38). Perhaps the greater immunoreactivity of LMP7 and increased expression of the LMP7 gene in RCC tissue observed in our study, compared to unchanged kidney tissue, may be associated with increased proteasomal processing of proteins in malignant cells.

In summary, our report shows significant alterations in the level and expression of the WNT/β-catenin pathway, CacyBP/SIP and LMP7 genes in human RCC tissues, which were dependent on the histological type of the tumor. The results of the study suggest an important role of the WNT/β-catenin pathway, CacyBP/SIP and LMP7 in RCC carcinogenesis and may indicate new aspects of pathomechanisms leading to differences in the biology of clear cell, papillary, and chromophobe RCC.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Bioethics Committee, Medical University of Bialystok (R-I-002/282/2019). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

IK and ZP: conceived of and designed the experiments. ZP, MA, MN, and GM: analyzed the data. IK, MA, MN, and GM: contributed reagents/materials/analysis tools. Writing—original draft preparation: ŻP. Writing—review and editing: IK. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by statutory funds from the Medical University of Bialystok.

REFERENCES

1. Muglia VF, Prando A. Renal cell carcinoma: histological classification and correlation with imaging findings. *Radiol Bras* (2015) 48:166–74. doi: 10.1590/0100-3984.2013.1927
2. Sadler GJ, Anderson MR, Moss MS, Wilson PG. Metastases from renal cell carcinoma presenting as gastrointestinal bleeding: two case reports and a review of the literature. *BMC Gastroenterol* (2007) 7:4. doi: 10.1186/1471-230X-7-4
3. Delahunt B, Bethwaite PB, Nacey JN. Outcome prediction for renal cell carcinoma: evaluation of prognostic factors for tumours divided according to histological subtype. *Pathology* (2007) 39:459–65. doi: 10.1080/00313020701570061
4. Gudbjartsson T, Hardarson S, Petursdottir V, Thoroddsen A, Magnusson J, Einarsson GV. Histological subtyping and nuclear grading of renal cell carcinoma and their implications for survival: a retrospective nation-wide study of 629 patients. *Eur Urol* (2005) 48:593–600. doi: 10.1016/j.eururo.2005.04.016
5. Polakis P. Wnt signaling in cancer. *Cold Spring Harb Perspect Biol* (2012) 4: pii: a008052. doi: 10.1101/cshperspect.a008052
6. Pulkkinen K, Murugan S, Vainio S. Wnt signaling in kidney development and disease. *Organogenesis* (2008) 4:55–9. doi: 10.4161/org.4.2.5849
7. Xu Q, Krause M, Samoylenko A, Vainio S. Wnt Signaling in Renal Cell Carcinoma. *Cancers (Basel)* (2016) 8:pii: E57. doi: 10.3390/cancers8060057

8. Majid S, Saini S, Dahiya R. Wnt signaling pathways in urological cancers: past decades and still growing. *Mol Cancer* (2012) 11:7. doi: 10.1186/1476-4598-11-7
9. Liu Z, Liu XW, Liu SA, Lv JJ, Fu Q. Clinical significance of changes of expression of the Wnt/β-catenin signaling pathway in renal clear cell carcinoma. *Eur Rev Med Pharmacol Sci* (2016) 20:4840–45.
10. Kruck S, Eyrich C, Scharpf M, Sievert KD, Fend F, Stenzl A, et al. Impact of an altered Wnt1/β-catenin expression on clinicopathology and prognosis in clear cell renal cell carcinoma. *Int J Mol Sci* (2013) 14:10944–57. doi: 10.3390/ijms140610944
11. Hsu RJ, Ho JY, Cha TL, Yu DS, Wu CL, Huang WP, et al. WNT10A plays an oncogenic role in renal cell carcinoma by activating WNT/β-catenin pathway. *PLoS One* (2012) 7:e47649. doi: 10.1371/journal.pone.0047649
12. Janssens N, Andries L, Janicot M, Perera T, Bakker A. Alteration of frizzled expression in renal cell carcinoma. *Tumour Biol* (2004) 25:161–71. doi: 10.1159/000081098
13. Bilim V, Ougolkov A, Yuuki K, Naito S, Kawazoe H, Muto A, et al. Glycogen synthase kinase-3: a new therapeutic target in renal cell carcinoma. *Br J Cancer* (2009) 101:2005–14. doi: 10.1038/sj.bjc.6605437
14. Kawazoe H, Bilim VN, Ugolkov AV, Yuuki K, Naito S, Nagaoka A, et al. GSK-3 inhibition in vitro and in vivo enhances antitumor effect of sorafenib in renal cell carcinoma (RCC). *Biochem Biophys Res Commun* (2012) 423(3):490–5. doi: 10.1016/j.bbrc.2012.05.147
15. Pal K, Cao Y, Gaisina IN, Bhattacharya S, Dutta SK, Wang E, et al. Inhibition of GSK-3 induces differentiation and impaired glucose metabolism in renal cancer. *Mol Cancer Ther* (2014) 13:285–96. doi: 10.1158/1535-7163.MCT-13-0681
16. Topolska-Woś AM, Chazin WJ, Filipek A. CacyBP/SIP – Structure and variety of functions. *Biochim Biophys Acta* (2016) 1860:79–85. doi: 10.1016/j.bbagen.2015.10.012
17. Zhai H, Shi Y, Jin H, Li Y, Lu Y, Chen X, et al. Expression of calcyclin-binding protein/Siah-1 interacting protein in normal and malignant human tissues: an immunohistochemical survey. *J Histochem Cytochem* (2008) 56:765–72. doi: 10.1369/jhc.2008.950519
18. Ning X, Chen Y, Wang X, Li Q, Sun S. The potential role of CacyBP/SIP in tumorigenesis. *Tumour Biol* (2016) 37:10785–91. doi: 10.1007/s13277-016-4871-y
19. Sun S, Ning X, Liu J, Liu L, Chen Y, Han S, et al. Overexpressed CacyBP/SIP leads to the suppression of growth in renal cell carcinoma. *Biochem Biophys Res Commun* (2007) 356:864–71. doi: 10.1016/j.bbrc.2007.03.080
20. Seliger B, Atkins D, Bock M, Ritz U, Ferrone S, Huber C, et al. Characterization of human lymphocyte antigen class I antigen-processing machinery defects in renal cell carcinoma lesions with special emphasis on transporter-associated with antigen-processing down-regulation. *Clin Cancer Res* (2003) 9:1721–7.
21. Seliger B, Höhne A, Knuth A, Bernhard H, Ehring B, Tampé R, et al. Reduced membrane major histocompatibility complex class I density and stability in a subset of human renal cell carcinomas with low TAP and LMP expression. *Clin Cancer Res* (1996) 2:1427–33.
22. Atkins D, Ferrone S, Schmahl GE, Störkel S, Seliger B. Down-regulation of HLA class I antigen processing molecules: an immune escape mechanism of renal cell carcinoma? *J Urol* (2004) 171:885–9. doi: 10.1097/01.ju.0000094807.95420.fe
23. Herman GE, Elfont EA. The taming of immunohistochemistry: the new era of quality control. *Biotech Histochem* (1991) 66(4):194–9. doi: 10.3109/10520299109109968
24. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
25. Znaor A, Lortet-Tieulent J, Laversanne M, Jemal A, Bray F. International variations and trends in renal cell carcinoma incidence and mortality. *Eur Urol* (2015) 67:519–30. doi: 10.1016/j.eururo.2014.10.002
26. Mehdi MZ, Nagi AH, Naseem N. MCM – 2 and Ki – 67 as proliferation markers in renal cell carcinoma: A quantitative and semi-quantitative analysis. *Int Braz J Urol* (2016) 42:1121–8. doi: 10.1590/S1677-5538.IBJU.2015.0388
27. Yang Q, Wang Y, Pan X, Ye J, Gan S, Qu F, et al. Frizzled 8 promotes the cell proliferation and metastasis of renal cell carcinoma. *Oncotarget* (2017) 8:78989–9002. doi: 10.18632/oncotarget.20742
28. von Schulz-Hausmann SA, Schmeel LC, Schmeel FC, Schmidt-Wolf IG. Targeting the Wnt/beta-catenin pathway in renal cell carcinoma. *Anticancer Res* (2014) 34:4101–8.
29. Koller CM, Kim Y, Schmidt-Wolf IG. Targeting renal cancer with a combination of WNT inhibitors and a bi-functional peptide. *Anticancer Res* (2013) 33:2435–40.
30. Sasamura H, Takahashi A, Yuan J, Kitamura H, Masumori N, Miyao N, et al. Antiproliferative and antiangiogenic activities of genistein in human renal cell carcinoma. *Urology* (2004) 64:389–93. doi: 10.1016/j.urology.2004.03.045
31. Chen S, Guttridge DC, You Z, Zhang Z, Fribley A, Mayo MW, et al. Wnt-1 signaling inhibits apoptosis by activating beta-catenin/T cell factor-mediated transcription. *J Cell Biol* (2001) 152:87–96. doi: 10.1083/jcb.152.1.87
32. Meyer T, Hart IR. Mechanisms of tumour metastasis. *Eur J Cancer* (1998) 34:214–21. doi: 10.1016/s0959-8049(97)10129-0
33. Langner C, Ratschek M, Rehak P, Schips L, Zigeuner R. Expression of MUC1 (EMA) and E-cadherin in renal cell carcinoma: a systematic immunohistochemical analysis of 188 cases. *Mod Pathol* (2004) 17:180–8. doi: 10.1038/modpathol.3800032
34. Kuehn A, Paner GP, Skinnider BF, Cohen C, Datta MW, Young AN, et al. Expression analysis of kidney-specific cadherin in a wide spectrum of traditional and newly recognized renal epithelial neoplasms: diagnostic and histogenetic implications. *Am J Surg Pathol* (2007) 31:1528–33. doi: 10.1097/PAS.0b013e318058818c
35. Shen SS, Krishna B, Chirala R, Amato RJ, Truong LD. Kidney-specific cadherin, a specific marker for the distal portion of the nephron and related renal neoplasms. *Mod Pathol* (2005) 18:933–40. doi: 10.1038/modpathol.3800373
36. Gerharz CD, Moll R, Störkel S, Ramp U, Thoenes W, Gabbert HE. Ultrastructural appearance and cytoskeletal architecture of the clear, chromophilic, and chromophobe types of human renal cell carcinoma in vitro. *Am J Pathol* (1993) 142:851–9.
37. Brembeck FH, Rosário M, Birchmeier W. Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. *Curr Opin Genet Dev* (2006) 16:51–9. doi: 10.1016/j.gde.2005.12.007
38. Frankland-Searby S, Bhaumik SR. The 26S proteasome complex: an attractive target for cancer therapy. *Biochim Biophys Acta* (2012) 1825:64–76. doi: 10.1016/j.bbcan.2011.10.003

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 © 2020 Piotrowska, Niezgoda, Młynarczyk, Acewicz and Kasacka. 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 2020
doi: 10.3389/fcell.2020.577784

Dissection of Anti-tumor Activity of Histone Deacetylase Inhibitor SAHA in Nasopharyngeal Carcinoma Cells *via* Quantitative Phosphoproteomics

Huichao Huang[1], Ying Fu[1], Ye Zhang[1], Fang Peng[1], Miaolong Lu[1], Yilu Feng[1], Lin Chen[2,3], Zhuchu Chen[1,4], Maoyu Li[1,4] and Yongheng Chen[1,5]**

[1] Department of Oncology, NHC Key Laboratory of Cancer Proteomics, XiangYa Hospital, Central South University, Changsha, China, [2] Molecular and Computational Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, CA, United States, [3] Department of Chemistry, University of Southern California, Los Angeles, CA, United States, [4] Department of Gastroenterology, XiangYa Hospital, Central South University, Changsha, China, [5] National Clinical Research Center for Geriatric Disorders, XiangYa Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Zexian Liu,
Sun Yat-sen University Cancer Center (SYSUCC), China
Ben Valdez,
University of Texas MD Anderson Cancer Center, United States

***Correspondence:**
Yongheng Chen
yonghenc@163.com
Maoyu Li
maoyuli@126.com

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

Received: 29 June 2020
Accepted: 27 October 2020
Published: 26 November 2020

Citation:
Huang H, Fu Y, Zhang Y, Peng F, Lu M, Feng Y, Chen L, Chen Z, Li M and Chen Y (2020) Dissection of Anti-tumor Activity of Histone Deacetylase Inhibitor SAHA in Nasopharyngeal Carcinoma Cells via Quantitative Phosphoproteomics. Front. Cell Dev. Biol. 8:577784. doi: 10.3389/fcell.2020.577784

Suberoylanilide hydroxamic acid (SAHA), a pan HDAC inhibitor, has been approved by the Food and Drug Administration (FDA) to treat cutaneous T cell lymphoma (CTCL). Nevertheless, the mechanisms underlying the therapeutic effects of SAHA on tumors are yet not fully understood. Protein phosphorylation is one of the most important means to regulate key biological processes (BPs), such as cell division, growth, migration, differentiation, and intercellular communication. Thus, investigation on the impacts of SAHA treatment on global cellular phosphorylation covering major signaling pathways deepens our understanding on its anti-tumor mechanisms. Here we comprehensively identified and quantified protein phosphorylation for the first time in nasopharyngeal carcinoma (NPC) cells upon SAHA treatment by combining tandem mass tags (TMTs)-based quantitative proteomics and titanium dioxide (TiO_2)-based phosphopeptide enrichment. In total, 7,430 phosphorylation sites on 2,456 phosphoproteins were identified in the NPC cell line 5-8F, of which 1,176 phosphorylation sites on 528 phosphoproteins were significantly elevated upon SAHA treatment. Gene ontology (GO) analysis showed that SAHA influenced several BPs, including mRNA/DNA processing and cell cycle. Furthermore, signaling pathway analysis and immunoblotting demonstrated that SAHA activated tumor suppressors like p53 and Rb1 *via* phosphorylation and promoted cell apoptosis in NPC cells but inactivated energetic pathways such as AMPK signaling. Overall, our study indicated that SAHA exerted anti-tumor roles in NPC cells, which may serve as novel therapeutic for NPC patients.

Keywords: histone deacetylase, suberoylanilide hydroxamic acid, p53–Rb1 signaling pathway, nasopharyngeal carcinoma, quantitative phosphoproteomic

Abbreviations: CCK-8, Cell Counting Kit 8; CTCL, cutaneous T cell lymphomas; DEPs, differentially expressed phosphoproteins; EBV, Epstein–Barr virus; FDA, Food and Drug Administration; HDACs, histone deacetylases; IMAC, immobilized metal affinity chromatography; LC–MS/MS, liquid chromatography–tandem mass spectrometry; NPC, nasopharyngeal carcinoma; PTMs, protein posttranslational modifications; SAHA, suberoylanilide hydroxamic acid; TiO_2, titanium dioxide.

INTRODUCTION

Nasopharyngeal carcinoma, as a common type of head and neck cancer, has a much higher incidence in Southern China and Southeast Asia, where the annual incidence is about 30 cases per 100,000 persons, in contrast to fewer than one case per 100,000 persons in the United States and Europe (Chang and Adami, 2006; Cao et al., 2011). EBV infection, genetic susceptibility, and environmental factors are involved in the development of NPC (Adham et al., 2012). Although progress has been made in the understanding, diagnosis, and treatment of NPC, it remains to be a serious threat to the community, urging for the development of early diagnostic biomarkers and novel therapeutics.

Histone deacetylases play key roles in regulating chromatin remodeling, the gene expressions of which they regulate epigenetically by turning down histone lysine acetylation in various pathophysiological conditions (Shahbazian and Grunstein, 2007). Both HDACs and histone lysine acetylation are involved in tumorigenesis, and inhibition of HDACs by specific inhibitors has emerged as an effective anti-tumor strategy (Wagner et al., 2010). SAHA (vorinostat), as an inhibitor of HDACs, has been approved by the FDA to treat refractory CTCL (Duvic et al., 2007; Marks and Breslow, 2007; Olsen et al., 2007). Besides CTCL, further research has demonstrated that SAHA processes anti-tumor effects in multiple types of solid tumors, including lung cancer, breast cancer, ovarian cancer, prostate cancer, neuroblastoma, as well as head and neck tumors (Munster et al., 2001; Komatsu et al., 2006; Konstantinopoulos et al., 2014; Xu et al., 2014; Wu et al., 2015). However, the anti-tumor role of SAHA has not been reported in NPC cells. In addition, previous studies merely focused on the alteration of protein acetylation to explain the anti-tumor mechanisms of SAHA.

Protein posttranslational modifications play critical roles in regulating fundamental biological functions, such as gene expression, signal transduction, and cell proliferation, division, and death (Blixt et al., 2004; Krueger and Srivastava, 2006; Hoffman et al., 2008). Prevalent PTMs include phosphorylation, acetylation, methylation, glycosylation, and ubiquitinylation. It has been well established that many PTMs contribute to abnormal cell proliferation, cell adhesion, and morphological changes during cancer progression (Sheehan et al., 2005; Krueger and Srivastava, 2006). Studies have shown that the dysregulated phosphorylation of some cellular signaling pathways contributes to oncogenesis, including receptor tyrosine kinases/PI3-kinase/Akt/mTOR, receptor tyrosine kinases/Ras/Raf/MEK/ERK, MEKK/MKK/JNK, and JAK/STAT (Shawver et al., 2002; Roberts and Der, 2007; Cicenas, 2008; Du et al., 2010). Meanwhile, phosphorylation and dephosphorylation are the key steps to regulate the activities of tumor suppressor genes p53 and Rb1, with the aberrant phosphorylation of both being strongly associated with cancer development, including NPC, breast cancer, lung cancer, and prostate cancer (Bischoff et al., 1990; Gubern et al., 2016; Sanidas et al., 2019). Therefore, the analysis of cancer phosphoproteome is crucial both for providing information on cancer cell signaling and for establishing the basis for targeted therapies.

Liquid chromatography–tandem mass spectrometry is the key technique to characterize and quantify phosphorylated sites and phosphorylated proteins. Due to the low abundance of phosphopeptides in the human body, there are four main strategies to enrich phosphopeptides: IMAC, TiO_2 affinity chromatography, strong cation exchange, and anti-phosphotyrosine antibodies. Currently, tandem mass tags (TMT)-based quantitative proteomics and TiO_2-based phosphopeptide enrichment coupled with LC–MS/MS are effective methods to detect, identify, and quantify phosphorylated sites and phosphorylated proteins.

In this study, we hypothesize that SAHA-mediated HDAC inhibition triggers globally reprogrammed signaling events *via* modulating the expression of signaling molecules, including kinases. Subsequently, protein phosphorylation alterations contribute to the inhibitory effect of SAHA on cancers. Combining quantitative phosphoproteomics and bioinformatic analysis, we demonstrated that SAHA treatment altered phosphorylation in key signaling pathways in NPC cells. Meanwhile, we indicated that SAHA inhibits proliferation and induces apoptosis in NPC cells. More importantly, our findings may provide a novel effective therapy for NPC patients and present a useful resource for future studies investigating the in-depth molecular mechanisms underlying the anti-tumor function of SAHA.

MATERIALS AND METHODS

Cell Culture and CCK-8 Assay

The cells were all purchased from American Type Culture Collection and cultured as previously described. 5-8F, HNE3, and 6-10B cells were cultured in RPMI-1640 medium containing 10% fetal bovine serum and 1% penicillin/streptomycin in a humidified environment at 37°C and 5% CO_2. Cell proliferation and cytotoxicity assay were performed by using CCK-8 (Dojindo, Kumamoto, Japan) according to the manufacturer's instructions. Briefly, NPC cells were seeded in 96-well plates in triplicate at an initial density of 5×10^3 cells/well. An increased concentration of SAHA (Sigma, St. Louis, MO, United States) was then added to each well for the indicated time. Subsequently, 10 μl of CCK-8 assay solution was added to each well, and the mixture was incubated for 2 h. The absorbance was measured at 450 nm using a multi-well spectrophotometer. Stock solutions of SAHA were prepared in dimethylsulfoxide (DMSO, Sigma–Aldrich) and diluted to the tested final concentrations in the culture medium. The final concentration of DMSO did not exceed 0.04%.

Colony Formation and Wound Healing Assay

Nasopharyngeal carcinoma cells were plated in six-well plates (1.0×10^3 cells per well) and cultured for 7 days. The cells were then fixed with paraformaldehyde for 10 min and stained with 1% crystal violet for 5 min prior to the counting of colonies. For the wound healing assay, NPC cells (2×10^5) were seeded in six-well plates. When confluency was reached, the cell monolayer was

scraped with a 10-μl pipette tip. Cell migration was observed by microscopy 24 h later.

Cell Apoptosis Assay

The apoptosis rate was evaluated using the annexin V-APC/7-AAD Apoptosis Detection kit (BD, 550474) according to the instructions from the manufacturer. Briefly, the cells were seeded into six-well tissue culture plates (4×10^5 cells/well). Following treatment, the cells were collected, washed twice with cold phosphate-buffered saline (PBS), and resuspended in 500 μl 1X binding buffer. Then, 5 μl annexin V-APC and 5 μl 7-AAD were added to the buffer and incubated at room temperature for 15 min in the dark. The cells were analyzed by flow cytometry (BD Biosciences, San Jose, CA, United States) within 1 h.

Western Blotting

Western blotting was performed as previously mentioned (Li et al., 2009). Briefly, NPC cells were lyzed as indicated in 0.3% Nonidet P40 (Sigma–Aldrich, 74388) buffer containing 150 mM NaCl, 50 mM Tris-HCl, pH = 7.5, and complete protease inhibitor cocktail (Roche, 04693132001). The following primary antibodies were commercially obtained: pan anti-acetyl-lysine (Kac) antibodies (PTM Biolab, 1:3,000 working dilution), H3 (Abcam, 18521; 1:1,000 working dilution), p-P53 Ser315 (Cell Signaling Technology, 2528; 1:1,000 working dilution), p-P53 ser37 (Cell Signaling Technology, 9289; 1:1,000 working dilution), P53 (Cell Signaling Technology, 9282; 1:1,000 working dilution), p-Rb1 ser807/811 (Cell Signaling Technology, 9308; 1:1,000 working dilution), Rb1 (Cell Signaling Technology, 9313; 1:1,000 working dilution), PARP 1 (Cell Signaling Technology, 9542; 1:1,000 working dilution), caspase 9 (Cell Signaling Technology, 52873; 1:1,000 working dilution), Bax (Cell Signaling Technology, 27745; 1:1,000 working dilution), Bcl2 (Cell Signaling Technology, 15071; 1:1,000 working dilution), ACTB (Sigma–Aldrich, A5441; 1:10,000 working dilution), ERK (Abcam, 17942; 1:1,000 working dilution), p-ERK Thr202/Tyr204 (Cell Signaling Technology, 9101; 1:1,000 working dilution), HER2 (Cell Signaling Technology, 2242; 1:1,000 working dilution), and p-HER2 Tyr1221/1222 (Cell Signaling Technology, 2249; 1:1,000 working dilution). ImageJ software (version 1.45s) was used to quantify the gray value of the western blot results. The western blot image was digitized to calculate the mean ± SD with Student's *t*-test ($p < 0.05$).

Histone Extraction

The core histones were extracted using a total histone extraction kit (The Epigentek Group Inc., OP-0006) according to the manufacturer's instructions. Briefly, NPC cells were harvested and re-suspended in 1× pre-lysis buffer, and the cells were lyzed on ice for 10 min with gentle stirring. Then, whole-cell lysates were further diluted in three volumes of lysis buffer and incubated on ice for 30 min. After that, the samples were centrifuged at 12,000 r/min for 5 min at 4°C, and the supernatant fraction (containing acid-soluble proteins) was transferred into a new vial with 0.3 volumes of the balance-DTT buffer added immediately.

Protein Extraction and Digestion

The cells were harvested, then washed with ice-cold PBS, and lyzed by incubation in SDS lysis buffer. After quantification of the total protein, protein digestion was performed according to the filter-aided sample preparation procedure. Briefly, 200 μg of proteins was reduced with 100 mM DTT, and then 200 μl UA buffer (8 M urea, 150 mM Tris-HCl, pH 8.0) was added. The mixture was then loaded into Microcon Ultracel YM-10 filtration device and centrifuged at 14,000 × *g* for 15 min. The concentrates were then diluted with 200 μl UA buffer and centrifuged at 14,000 × *g* for 15 min. After centrifugation, the concentrates were alkylated in 100 μl IAA (50 mM IAA in UA) for 30 min in the dark. After centrifugation for 10 min, the concentrates were washed twice with UA buffer and twice with 100 mM NH_4HCO_3. Subsequently, trypsin solution (8 μg trypsin in 40 μl NH_4HCO_3 buffer) was added to the filter, and the proteins were incubated at 37°C overnight. Tryptic peptides were collected by centrifugation, followed by an additional wash with an elution solution [70% acetonitrile (ACN) and 0.1% formic acid]. Finally, the peptide mixture was desalted with C18-SD Extraction Disk Cartridge, and the peptide concentration was assayed by measuring the absorbance at 280 nm.

Tandem Mass Tagging Labeling

Experiments were performed for three biological replicates. For phosphoproteome, three technical replicates were applied as well. The dried peptides were redissolved in 0.5 M TEAB and processed according to the manufacturer's protocol for a six-plex TMT kit (Thermo Fisher Scientific). Briefly, one unit of TMT reagent (used to label the tryptic peptides of 100 μg proteins) was thawed and reconstituted in 24 μl ACN. The peptide mixture was incubated with the prepared TMT reagent (1 h, room temperature) and then quenched by the addition of 200 μl of 5% hydroxylamine for 15 min. Next, the TMT-labeled peptide mixtures were pooled equally (1:1:1:1:1:1), desalted, and dried by vacuum centrifugation.

Phosphopeptide Enrichment

Phosphopeptide enrichment was performed as described by Larsen et al. (2005). Briefly, lyophilized peptides were re-suspended in DHB buffer [3% w/v DHB, 80% v/v ACN, 0.1% v/v trifluoroacetic acid (TFA)]. Then, TiO_2 beads (GL Sciences, Japan) were added, and the mixture was agitated for 40 min. TiO_2 beads were recovered by centrifugation and washed three times with washing buffer I (30% ACN and 3% TFA) and three times with washing buffer II (80% ACN and 0.3% TFA). Lastly, the phosphopeptides were eluted with the elution buffer (5% NH_4OH/50% ACN), followed by lyophilization and MS analysis.

LC–ESI–MS/MS Analysis by Q-Extractive MS

Peptides were dissolved in solvent A (0.1% FA) and loaded onto a Thermo scientific EASY column (C18 column, 5 μm, 100 μm × 2 cm, Thermo Scientific). Peptide separation was performed using a reversed-phase analytical column (C18 column, 75 μm × 250 mm, 3 μm, Thermo Scientific). The

gradient was comprised of an increase from 0 to 55% solvent B (0.1% FA in 98% ACN) for 220 min, 55 to 100% for 8 min, and then holding at 100% for the last 12 min, at a constant flow rate of 250 nl/min on an EASY-nLC 1000 UPLC system. The eluted peptides were analyzed by Q Exactive^{TM} hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific). A data-dependent procedure was one MS scan (m/z range of 350–1,800) followed by 10 MS/MS scans for the top 20 precursor ions. Dynamic exclusion was enabled, with exclusion duration of 30 s. Automatic gain control was set as 3e6 to prevent overfilling of the ion trap. The peptides were detected in the Orbitrap at a resolution of 70,000. Peptides were selected for MS/MS using NCE setting as 29, and ion fragments were detected at a resolution of 17,500.

MS Data Analysis

MS/MS spectra were searched using Mascot 2.2 (Matrix Science) embedded in Proteome Discoverer 1.4 against the UniProt human FASTA (released on 5/5/2018). For protein identification, the following options were used: peptide mass tolerance, 20 ppm; MS/MS tolerance, 0.1 Da; enzyme, trypsin; missed cleavage, 2; fixed modifications, carbamidomethyl (C); variable modifications, TMT 6-plex (N-term), TMT 6-plex (K), oxidation (M), phosphorylation (S/T/Y), and false discovery rate (FDR) $\leq$ 0.01. Proteome Discoverer 1.4 was used to extract the peak intensity of each expected TMT reporter ion from the fragmentation spectrum. Only spectra in which all quant channels are present were used for quantification. The score threshold for peptide identification was set at 1% FDR and with PhosphoRS site probability cutoff of 0.75. Student's t-test was used to evaluate the statistical significance, and FDR was calculated. The criteria for significant abundance changes were abundance ratios $\geq$ 1.2 and P-value $\leq$ 0.05.

Bioinformatic Analysis

Metascape, a gene annotation and analysis resource[1], was employed to conduct gene enrichment and functional annotation analyses. The gene ontology (GO) annotation of DEPs was derived from the UniProt-GOA database[2]. DEPs were classified by GO annotation based on three categories, including biological processes (BPs), cellular compartments, and molecular functions (MFs), and $p < 0.05$ was considered statistically significant. Pathway enrichment analysis was performed by IPA software (Ingenuity Systems) and KEGG pathway database[3]. The adjusted P-values < 0.05 were considered statistically significant.

Statistical Analysis

Two-tailed Student's t-tests were used for all comparisons. All values included in the figures represent the mean values $\pm$ SD. Error bars represent $\pm$ SD for triplicate experiments. The statistical significance is indicated with asterisks (*). A two-sided P-value < 0.05 was considered statistically significant ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

[1]http://metascape.org

[2]http://www.ebi.ac.uk/GOA/

[3]http://kobas.cbi.pku.edu.cn/

RESULTS

SAHA Inhibits the Proliferation of NPC Cells

To investigate the effects of SAHA on NPC cell proliferation, NPC cells (5-8F, 6-10B, and HNE3) were treated with increasing concentrations of SAHA. SAHA treatment markedly decreased cell survival compared with the untreated controls in a dose- and time-dependent manner (**Figure 1A** and **Supplementary Figure S1**). Similarly, colony formations were robustly inhibited in HNE3 and 5-8F cells upon SAHA treatment (**Figure 1B**). Next, we performed wound healing assay and observed that SAHA treatment intensively suppressed cell migration in NPC cells (**Figures 1C,D**). As SAHA is a pan HDAC inhibitor, we examined the lysine acetylation level of histone H3 in NPC cells by western blotting. As shown in **Figure 1E**, a significant augmentation of H3Kac signals was detected as the concentration and the treatment time of SAHA increased. The maximum H3Kac signal was detected at 6 μM and 24 h of SAHA treatment. Since cell viability was still nearly 80% when treated at 6 μM for 24 h, this condition was applied for all the following experiments in NPC cells. Interestingly, we observed that SAHA altered HER2 and ERK phosphorylation in 5-8F cells utilizing the handy phosphorylated antibodies (**Figure 1F**), which gave us hints to investigate protein phosphorylation upon SAHA stimulation.

Taken together, these results showed that SAHA inhibited the proliferation and the migration of NPC cells. Meanwhile, protein phosphorylation was regulated by SAHA treatment.

Profile of Phosphoproteomics Data Toward SAHA Treatment in NPC Cells

In order to better understand the anti-tumor mechanisms of SAHA in NPC cells, 5-8F cells were treated with SAHA or DMSO, and mass spectrometry was performed on whole-cell lysates with quantitative phosphoproteomic technology by six-plex TMT to detect alterations in protein phosphorylation. The workflow of TMT-based phosphoproteomic analysis is demonstrated in **Figure 2**.

Overall, 7,430 phosphorylation sites in 2,424 phosphoproteins were identified, of which 4,318 phosphorylation sites in 1,836 phosphoproteins were quantified (**Supplementary Table 1**). A total of 6,662 (89.66%) of the sites were found at serine, 743 (10%) at threonine, and 25 (0.34%) at tyrosine residues (**Figure 3A**). Among the quantified 4,318 phosphopeptides, 1,898 (43.96%) phosphopeptides contained single phosphorylation, 1,868 (43.26%) phosphopeptides had two phosphorylated sites, and 552 (12.78%) phosphopeptides were phosphorylated at more than two sites (**Figure 3B**). For proteins, 902 (49.13%) of 1,836 phosphoproteins contained a single phosphorylated site, 399 (21.73%) proteins harbored two phosphorylated sites, and 535 (29.14%) proteins have three or more than three phosphorylated sites (**Figure 3C**).

Among these phosphorylation sites, 2,098 showed a significant change upon SAHA stimulation ($p < 0.05$; fold change > 1.2 or < -1.2). A total of 847 phosphorylation sites were

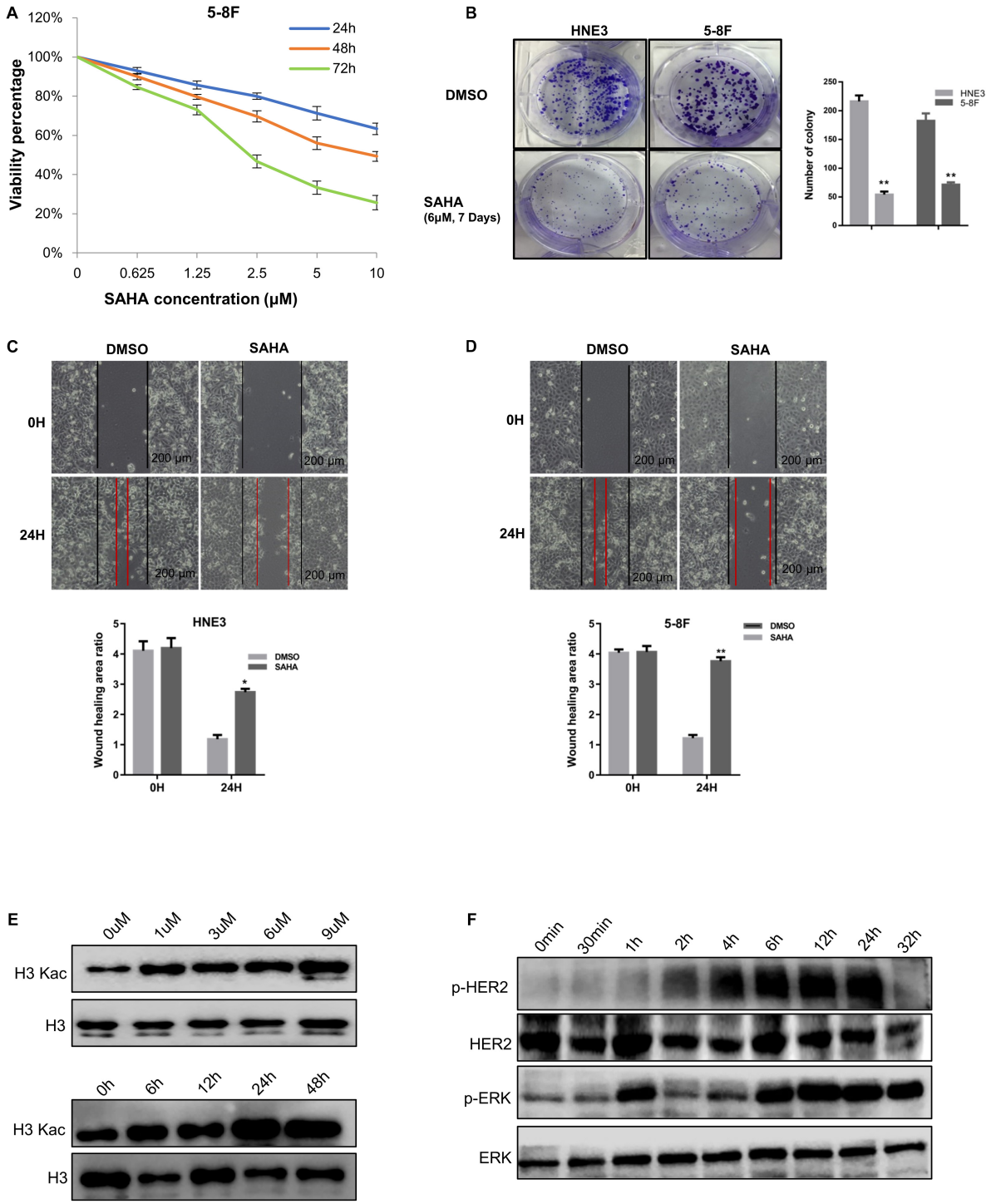


FIGURE 1 | Suberoylanilide hydroxamic acid (SAHA) inhibits cell proliferation and migration in nasopharyngeal carcinoma (NPC) cells. **(A)** SAHA treatment suppressed NPC cell growth. 5-8F cells were treated with various concentrations of SAHA for 24, 48, and 72 h, respectively. **(B)** SAHA treatment inhibited the proliferation of NPC cells. HNE3 and 5-8F cells were treated with SAHA or dimethylsulfoxide (DMSO) as indicated and analyzed by a colony formation assay (left panel). A quantitative analysis of the colony was performed by ImageJ (right panel). **(C,D)** SAHA treatment inhibited the migration of NPC cells. HNE3 **(C)** and 5-8F cells **(D)** were treated with SAHA or DMSO as indicated and analyzed by wound healing assay. Scale bars: 200 μm. A quantitative analysis of the wound healing area was performed by ImageJ. **(E)** SAHA treatment affected the acetylation of histone H3 in NPC cells. 5-8F cells were treated with DMSO or SAHA at the indicated concentrations (upper panel) and time points (lower panel). **(F)** SAHA treatment affected protein phosphorylation in NPC cells. Immunoblots of whole-cell lysates from 5-8F cells treated with DMSO or SAHA at indicated time points were analyzed. Data are shown as mean ± SD ($n = 3$) or typical photographs of one representative experiment. Similar results were obtained in three independent experiments. $^{*}p < 0.05$, $^{**}p < 0.01$.

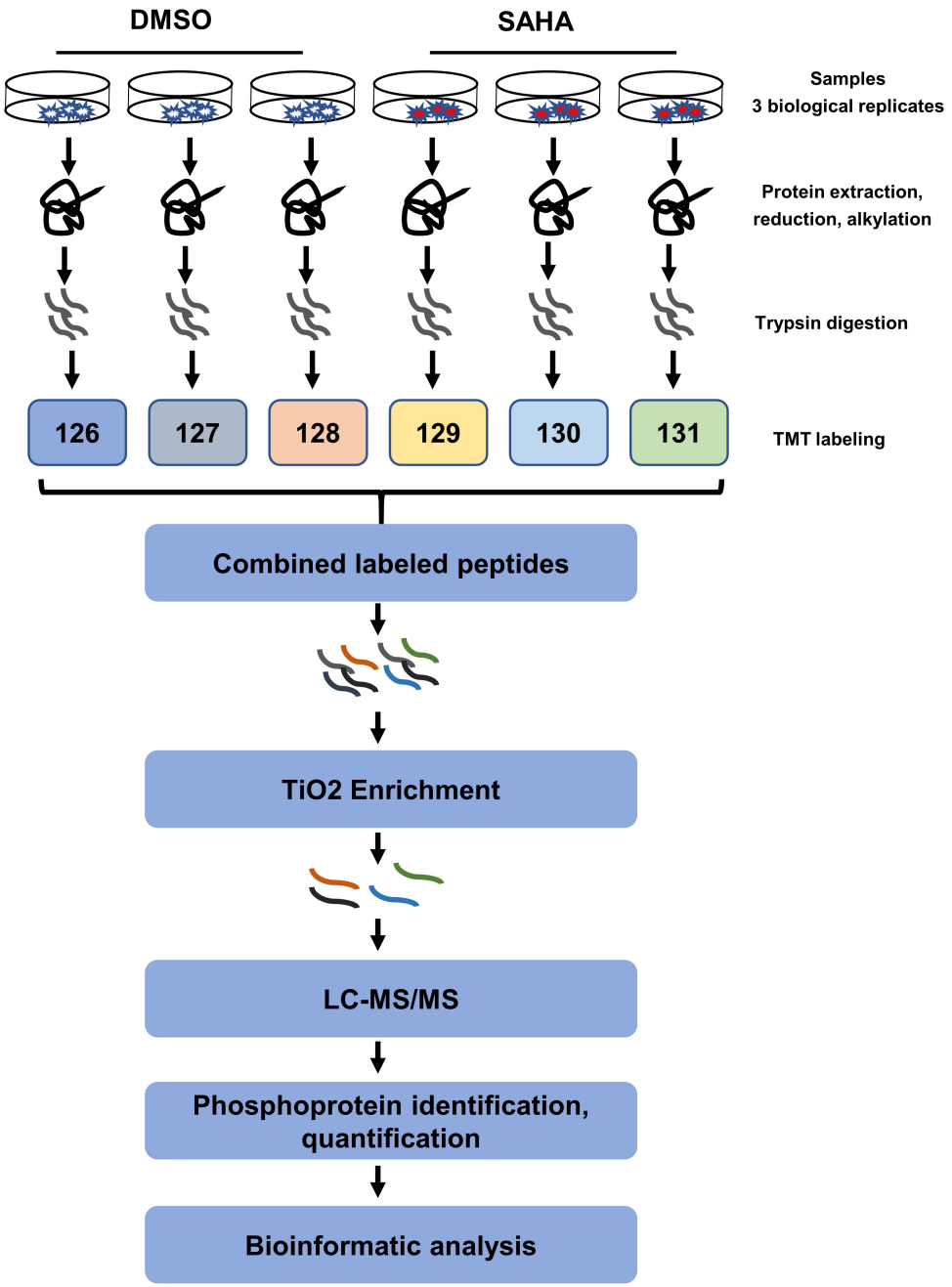


FIGURE 2 | Schematic illustration of the (TMT)-based quantitative phosphoproteomic workflow. 5-8F cells treated with suberoylanilide hydroxamic acid for 24 h were subjected to six-plex TMT labeling. Combined labeled peptides were subjected to TiO_2 enrichment. The enriched phosphopeptides were analyzed using an Orbitrap-equipped mass spectrometer. TMT, tandem mass tag; LC–MS/MS, liquid chromatography–tandem mass spectrometry; TiO_2, titanium dioxide.

upregulated and 647 sites were downregulated upon SAHA treatment (**Figure 3D**).

Identification of Differentially Expressed Phosphoproteins

We first evaluated the reproducibility of the quantification. Pearson correlation coefficient analysis for phosphopeptides based on the ratio was performed to assess the similarity between biological replicates (**Figure 4A**). A higher correlation between any two of the three replicates was observed with R^2 value higher than or equal to 0.82, indicating that the results had reasonable technical and instrument variation.

To determine the involved protein categories, functions, and localizations of the DEPs upon SAHA treatment, the GO-based classification and enrichment analysis were conducted.

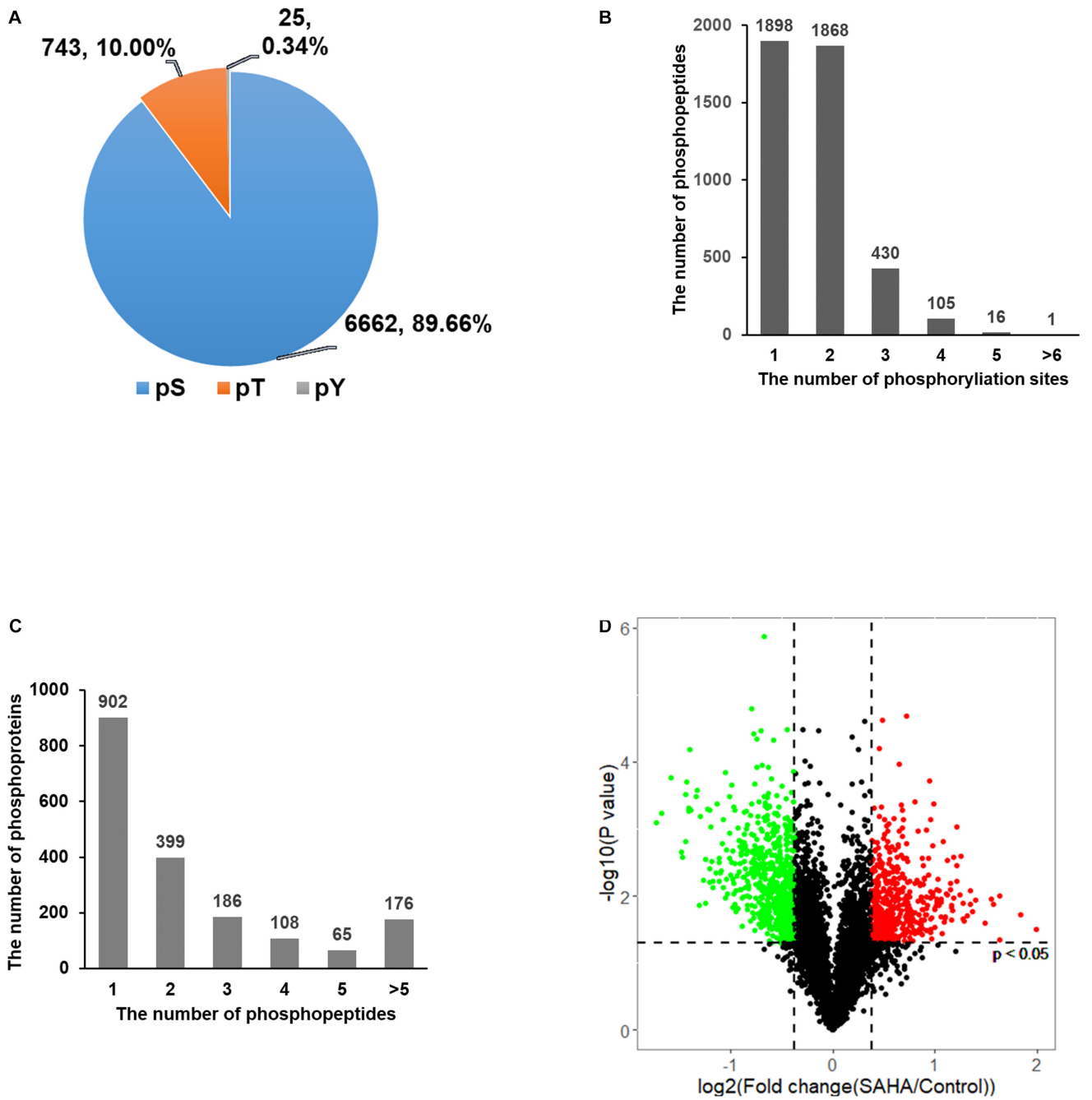


FIGURE 3 | Distribution of phosphorylation sites. **(A)** The distribution of phosphorylation on serine, threonine, and tyrosine showed that phosphorylated serine was dominant. **(B)** The distribution of peptides with single- and multi-phosphorylation sites showed that the majority of phosphopeptides have only one or two phosphorylation sites. **(C)** The distribution of phosphoproteins based on the number of identified phosphopeptides showed that nearly half of the proteins have only one phosphopeptide. **(D)** Volcano plots of group comparisons [suberoylanilide hydroxamic acid (SAHA) *versus* dimethylsulfoxide] showing the adjusted significance *P*-value (log2) *versus* fold change (log2). The plots indicate the most robust protein changes in SAHA. The horizontal gray dotted lines indicate an adjusted *P*-value threshold of 0.05; the vertical gray dotted lines indicate a fold change threshold of 20%. Significantly upregulated phosphopeptides are marked in red and those downregulated are in green (adjusted *p*-value < 0.05).

As shown in **Figure 4B**, the MF category indicated that the DEPs are mainly associated with enzymes (464, 53%), peptidases (140, 16%), transcription regulators (136, 16%), and kinases (55, 6%). The cellular component category showed that DEPs are mainly involved in the nucleus (438, 50%), cytoplasm (293, 33%), plasma membrane (75, 9%), and extracellular space (17, 2%) (**Figure 4C**). In BP analysis, DEPs were markedly enriched in cellular component organization or biogenesis (467, 53%),

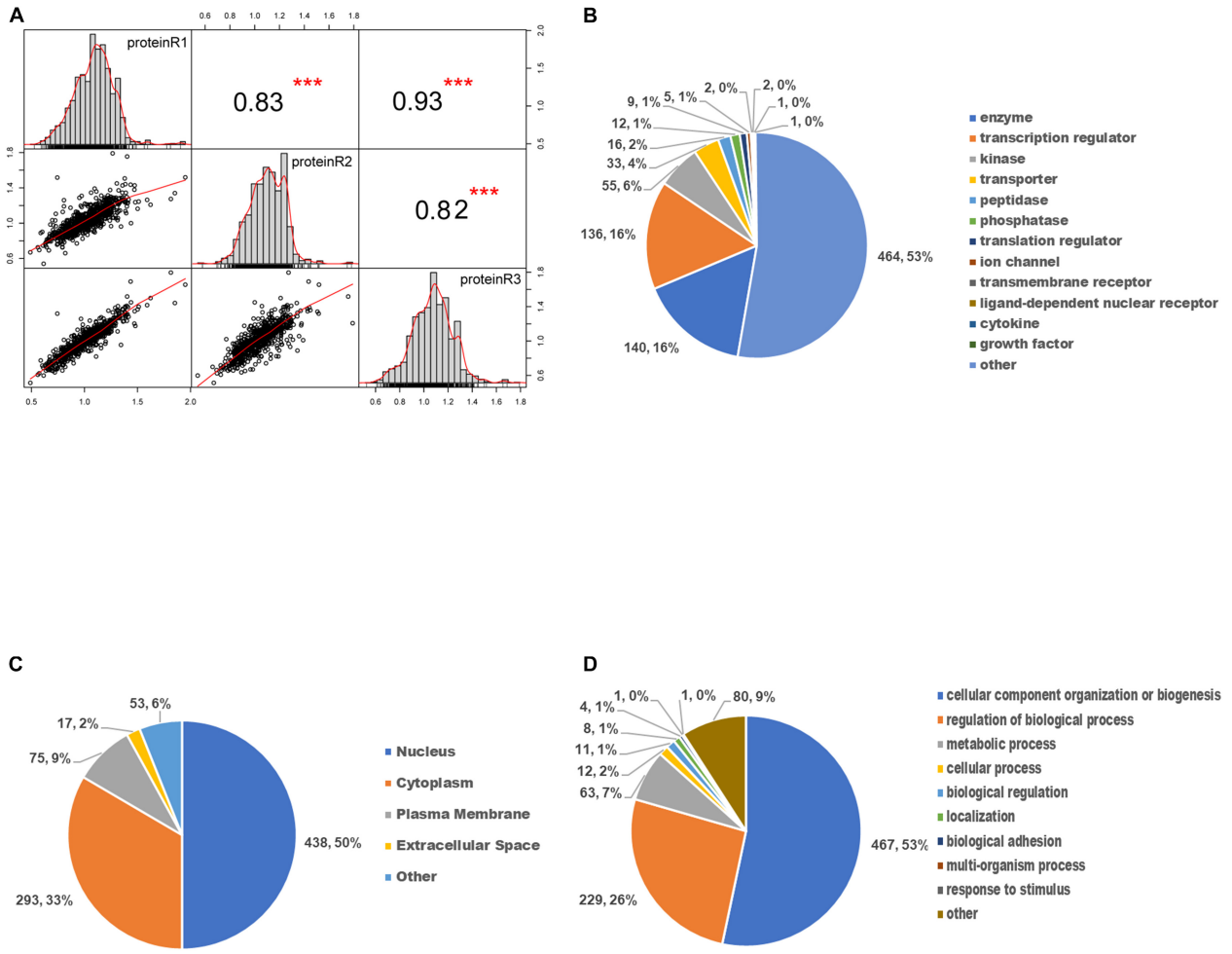


FIGURE 4 | Gene ontology enrichment analysis of differentially expressed phosphoproteins. **(A)** The scatter plot matrix of Pearson correlation between replicate experiments (proteinR1-R3) shows that the reproducibility is high (***$p < 0.001$). **(B)** Molecular function analysis of differentially expressed phosphoproteins. **(C)** Cellular component analysis of differentially expressed phosphoproteins. **(D)** Biological process analysis of differentially expressed phosphoproteins.

regulation of BP (229, 26%), metabolic process (63, 7%), and cellular process (12, 2%) (**Figure 4D**).

Functional Characteristics and IPA Pathway Analysis of DEPs

To obtain a global view of the cellular functions of the DEPs toward SAHA treatment, we performed comprehensive bioinformatic analysis based on GO, KEGG pathway, and protein complex database (**Figure 5**). According to their membership similarities, enriched processes or pathways were grouped into different clusters. In **Figure 5A**, the top 20 enrichment processes or pathways are presented, with different notes referring to their corresponding enriched terms. The colors represented different cluster IDs, of which nodes of the same color indicated the similarity of their cellular functions. Then, to further study these enriched processes, we ranked them based on the enrichment degree (**Figure 5B**). The results showed that the processes or pathways associated with mRNA processing, cell cycle, cell division, ribonucleoprotein complex biogenesis, covalent chromatin modification, and nuclear transport were robustly enriched, which might serve as the potential targets of SAHA.

To identify cellular signaling associated with the anti-tumor effects of SAHA in NPC cells, IPA signaling pathway analysis was performed. Overall, the DEPs were involved in 32 significant cellular pathways ($p < 0.05$ and FDR < 0.05) (**Figure 6A** and **Supplementary Table 3**). A comprehensive analysis of these pathways revealed that DEPs were markedly enriched in five important tumor-related molecular network systems, including cell proliferation, cell metabolism, cell cycle, cell adhesion, and cell signal transduction (**Table 1**). Upon SAHA treatment, the upregulated DEPs were mainly enriched in cell cycle, cell apoptosis, and DNA repair related pathways, while the downregulated DEPs were mainly harbored in energy metabolism, signal transduction, and cell proliferation relevant pathways (**Figure 6A**).

Among the activated pathways found in this study, p53 signaling pathway, G2/M DNA damage checkpoint regulation,

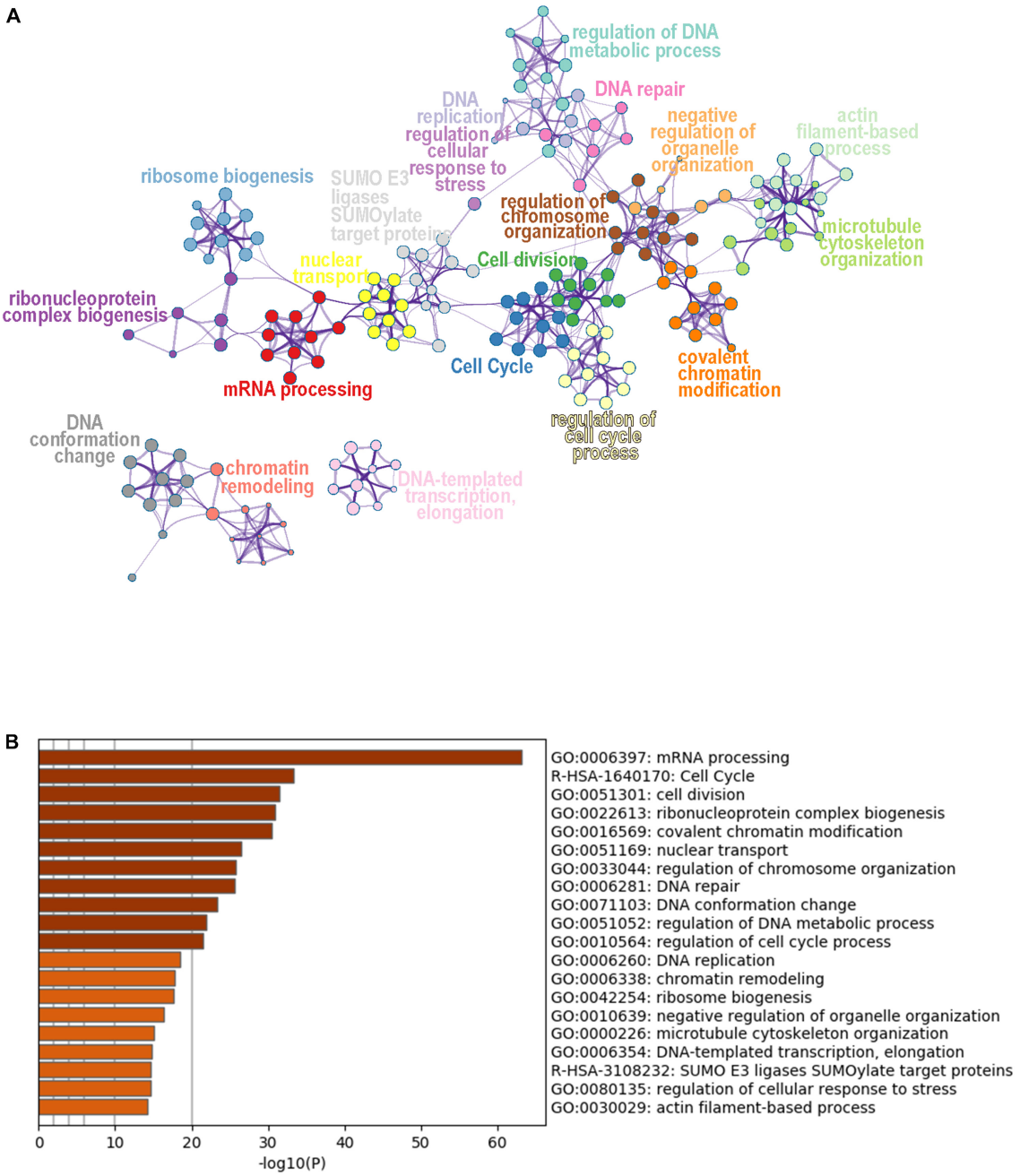


FIGURE 5 | Cluster analysis of differentially expressed phosphoproteins. **(A)** Network plot of the relationships among enriched terms. Nodes representing enriched terms were grouped into clusters based on their membership similarities and colored by their cluster ID; nodes that share the same cluster ID are typically close to each other. **(B)** Bar graph of the top 20 enriched terms across input gene lists, colored by *p*-value.

and the role of BRCA1 in DNA damage response were the three most significantly activated (z-score > 2) in response to SAHA treatment (**Figure 6A**). Representative p53 signaling pathways that are involved in cell cycle, cell proliferation, cell apoptosis, and other critical cellular processes are among the most robustly activated pathways (**Figure 6B**). In this study, 10 DEPs (CCNK, CDK2, CHEK1, CSNK1D, HDAC1, MED1, PIK3C2A, RB1, TOPBP1, and TP53) were identified in p53 pathway (**Table 2** and **Supplementary Table 4**). On the other hand, the pathways associated with energy metabolism, including AMPK pathway, IGF-1 pathway, insulin receptor pathway, and PFKFB4 signaling pathway, were inhibited by SAHA, with AMPK pathway being the most inhibited one. A total of 12 DEPs (PPM1H, PPP2R5E, PRKAA1, PRKACA, PRKAR1A, PRKAR1B, RAB1A, RAB8A, RPS6KB1, SMARCA4, SMARCC1, and ULK1) involved in AMPK pathway were discovered *via* IPA signaling analysis

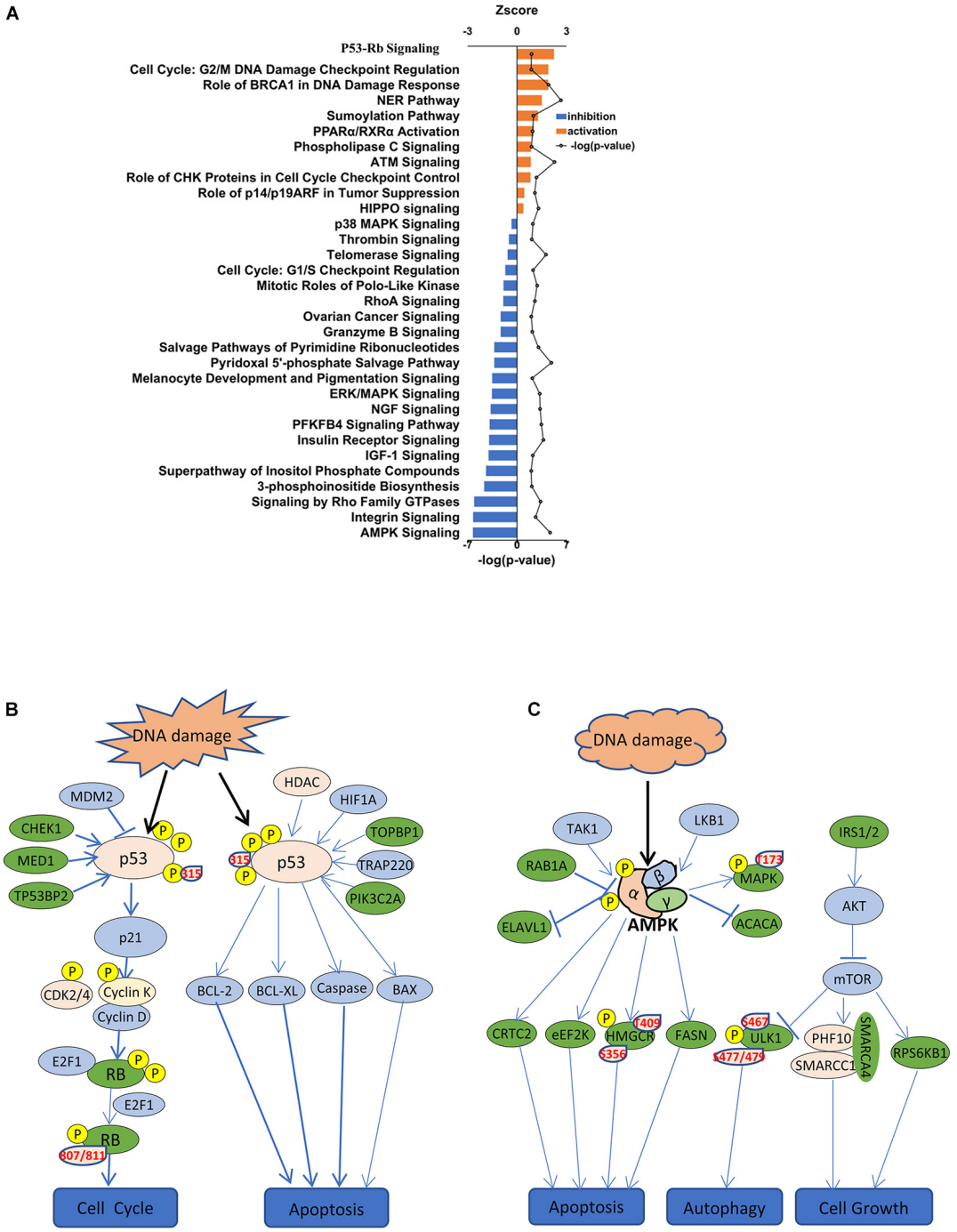


FIGURE 6 | IPA signaling pathway analysis of differentially expressed phosphoproteins. **(A)** IPA showing decreased and increased biological functions of differentially expressed phosphoproteins in suberoylanilide hydroxamic acid-treated cells. Depicted are functions with an activation score (*z*-score) > 0 (increased activation) or < 0 (decreased activation). **(B)** Activated p53 signaling pathway is presented. Significantly upregulated phosphoproteins are shown in pink, while downregulated phosphoproteins are shown in green. **(C)** Inhibited AMPK signaling pathway is presented. Significantly upregulated phosphoproteins are shown in pink, while downregulated phosphoproteins are shown in green.

(**Table 3** and **Supplementary Table 5**), with several representative proteins presented in **Figure 6C**. The raw data of p53 signaling and AMPK signaling pathway from IPA database, along with information on all the identified proteins, are summarized in **Supplementary Tables 4, 5**. Collectively, these results revealed that SAHA manipulated the activity of critical molecules *via* phosphorylation to activate P53 and inactivate AMPK pathways, leading to the proliferation arrest and apoptosis of NPC cells.

TABLE 1 | The functional categories of identified pathways *via* IPA.

Category	Pathway	z-Score
Cell proliferation		
	P53 signaling	2.236
	HIPPO signaling	0.378
	Role of p14/p19ARF in tumor suppression	0.447
	PPARA/RXRA activation	1
	Granzyme B signaling	−1
	ERK/MAPK pathway	−1.528
	NGF signaling	−1.604
	3-Phosphoinositide biosynthesis	−2
Cell cycle		
	Cell cycle: G2/M DNA damage checkpoint regulation	1.89
	ATM signaling	0.832
	Role of CHK proteins in cell cycle checkpoint control	0.816
	Role of BRCA1 in DNA damage response	1.89
	Cell cycle G1/S checkpoint regulation	−0.707
	P38 MAPK pathway	−0.333
	Mitotic roles of Polo-like kinases	−0.816
	Melanocyte development and pigmentation signaling	−1.508
	Integrin signaling	−2.668
Cell metabolism		
	AMPK signaling	−2.683
	NER pathway	1.5
	Salvage pathways of pyrimidine ribonucleotides	−1.387
	Pyridoxal 5′-phosphate salvage pathway	−1.387
	PFKFB4 signaling	−1.667
	Insulin receptor signaling	−1.698
	IGF-1 signaling	−1.732
	Superpathway of inositol phosphate compounds	−1.886
Cell adhesion		
	Thrombin signaling	−0.5
	Telomerase signaling	−0.577
	RhoA signaling	−0.832
	Ovarian cancer signaling	−1
	Signaling by Rho family GTPases	−2.6
Cell signal transduction		
	Sumoylation pathway	1.265
	Phospholipase C signaling	0.853

Validating the Activation of p53–Rb1 Axis by SAHA Treatment in NPC Cells

As presented in **Figure 6B**, our quantitative phosphoproteomic data indicated that p53 phosphorylation at Ser315 residue was increased upon SAHA treatment in NPC cells. To validate this finding, we assessed the phosphorylation level of p53 at Ser315 site as well as at Ser37 site by western blotting analysis. As shown in **Figures 7A,B**, the phosphorylation signals of Ser315 residue, but not Ser37 residue, were significantly elevated by SAHA treatment in different NPC cell lines, including 5-8F, HNE3, and 6-10B cells ($P < 0.001$).

Rb1 is a well-known tumor suppressor gene, the dephosphorylated type of which binds E2F and blocks its ability to activate transcription. Here we observed that several sites (S807/S811, T821/T826, S788/S795, S37, T821/T826, S612, and S249/T252) were downregulated on the phosphorylation level of Rb1 upon SAHA treatment, of which S807/S811 is the most downregulated (**Table 2**). A representative MS2 spectrum of Rb1 phosphorylated peptide containing two serine phosphorylation sites is presented (**Figure 7C**). Consistent with the quantitative phosphoproteomic analysis, the immunoblot results demonstrated that the phosphorylation signals of S807/S811 in Rb1 were reduced in SAHA-treated cells (**Figures 7A,B**). These results suggested that SAHA may exert its anti-tumor effect by regulating the phosphorylation of p53 and Rb1 at specific residues.

Since cell apoptosis-related pathways were markedly enriched upon SAHA treatment, we next determined the expression of PARP, Bax, and Bcl2 in NPC cells with or without SAHA treatment by western blotting. Compared with DMSO-treated cells, the expression levels of cleaved PARP 1, cleaved caspase 9, and Bax, but not Bcl2, were significantly increased by SAHA treatment (**Figure 7D**). The quantitative analyses of these proteins normalized to ACTB are shown in **Figure 7E**. This potentiation of apoptosis was further confirmed by flow cytometry analysis (**Figures 7F,G**). These observations demonstrated that SAHA treatment promoted cell apoptosis in NPC cells.

Taken together, our data indicated that SAHA activates tumor suppressors like p53 and Rb1 *via* phosphorylation and promotes cell apoptosis in NPC cells to implement its anti-tumor role.

DISCUSSION

Nasopharyngeal carcinoma is a highly malignant head and neck tumor. Despite that significant efforts have been put into investigation, early diagnostic biomarkers and efficient therapies are lacking (Cao et al., 2011), bringing a huge burden to the public health community and urging for novel therapies for NPC patients. Here we found that SAHA, a clinically approved drug, inhibited the proliferation and promoted the apoptosis of NPC cells. Applying quantitative phosphoproteomic analysis, we reported for the first time the phosphoproteomic profile and deciphered the potential anti-tumor mechanism of SAHA in NPC cells. Overall, our findings provided novel insights into the anti-tumor effect of SAHA in NPC cells, which may lead to new therapies for NPC patients.

As the most promising deacetylation inhibitor, SAHA has exhibited potent anti-tumor properties both in solid tumors and hematologic malignances *via* diverse mechanisms (Minucci and Pelicci, 2006; Siegel et al., 2009). In this study, we applied TMT-based quantitative proteomics associated with TiO_2-based

TABLE 2 | Description of identified molecules in the P53 pathway.

Accession	Symbol	Description	Phosphorylation sites	*p*-Value
G3V5E1	CCNK	Cyclin K	S340	0.0322
G3V317	CDK2	Cyclin-dependent kinase 2	Y15	0.000411
J3KN87	CHEK1	Checkpoint kinase 1	S331	0.0296
B4DIJ9	CSNK1D	Casein kinase 1 delta ase	T236	0.0132
Q5HYD4	HDAC1	Histone diacetyl 1	S421/S423	0.00595
B2RAG9	MED1	Mediator complex subunit 1	T1051/T1057	0.0144
A0A0C4DGF9	PIK3C2A	Phosphatidylinositol-4-phosphate3-kinase catalytic subunit type2 alpha	S259	0.0338
P06400	RB1	RB transcriptional corepressor 1	S807/S811, T821/T826, S788/S795, S37, T821/T826, S612, S249/T252	0.0254
A0AV47	TOPBP1	DNA topoisomerase II binding protein 1	T774/S777	0.0141
I3L0W9	TP53	Tumor protein p53	S315	0.00857
B4DI25	TP53BP2	Tumor protein p53 binding protein 2	S183/S196	0.0163

TABLE 3 | Description of identified molecules in the AMPK pathway.

Accession	Symbol	Description	Phosphorylation sites	*p*-Value
Q13085	ACACA	Acetyl-CoA carboxylase alpha	S2343	0.00832
Q4LE49	ARID1A	AT-rich interaction domain 1A	S418/S426	0.000503
Q5T4K5	CRTC2	CREB-regulated transcription coactivator 2	S136/T138	0.0284
O00418	EEF2K	Eukaryotic elongation factor 2 kinase	S71	0.0113
Q15717	ELAVL1	ELAV-like RNA binding protein 1	S202	0.00685
A0A0U1RQF0	FASN	Fatty acid synthase	S207	0.00722
A8KA27	HMGCR	3-Hydroxy-3-methylglutaryl-CoA reductase	T409, S356	0.0154
P35568	IRS1	Insulin receptor substrate 1	S629/S636	0.0034
Q96RG4	IRS2	Insulin receptor substrate 2	S736/S737	0.0212
J3QL77	MAP2K3	Mitogen-activated protein kinase 3	S3/S15	0.0357
Q9UG54	MAP3K7	Mitogen-activated protein kinase 7	S93	0.0485
B4DHN0	MAPK1	Mitogen-activated protein kinase 1	T173	0.0178
B4DK20	PHF10	PHD finger protein10	S115/S119	0.00465
A0A0C4DGF9	PIK3C2A	Phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 alpha	S259	0.0338
Q9ULR3	PPM1H	Protein phosphatase, $Mg_2 + Mn_2 +$ dependent 1H	S221/S223	0.0188
Q86XZ2	PPP2R5E	Protein phosphatase 2 regulatory subunit B'epsilon	S33/S34	0.0465
Q13131	PRKAA1	Protein kinase AMP-activated catalytic subunit alpha 1	S486/T490	0.0259
Q15136	PRKACA	Protein kinase cAMP-activated catalytic subunit alpha	T179	0.00915
K7EMU2	PRKAR1A	Protein kinase cAMP-activated catalytic subunit alpha	S83	0.00108
C9J4C2	PRKAR1B	Protein kinase cAMP-dependent type I regulatory subunit beta	S77/S83, S71/S77/S83	0.00282
Q96RD8	RAB1A	RAB1A, member RAS oncogene family	T84	0.0223
P61006	RAB8A	RAB8A, member RAS oncogene family	S181/S185	0.00109
B4DDM0	RPS6KB1	Ribosomal protein S6 kinase B1	S403/S408, S397/T400/S403/S408	0.00092
P51532	SMARCA4	SWISNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, member 4	S613, S1417/T1423, T1423S695/S699	0.00768
Q05BW5	SMARCC1	SWISNF-related, matrix-associated, actin-dependent regulator of chromatin subfamily c member 1	S328/S330, S310	0.00936
O75385	ULK1	unc-51 like autophagy activating kinase 1	S467, S477/S479	0.0137

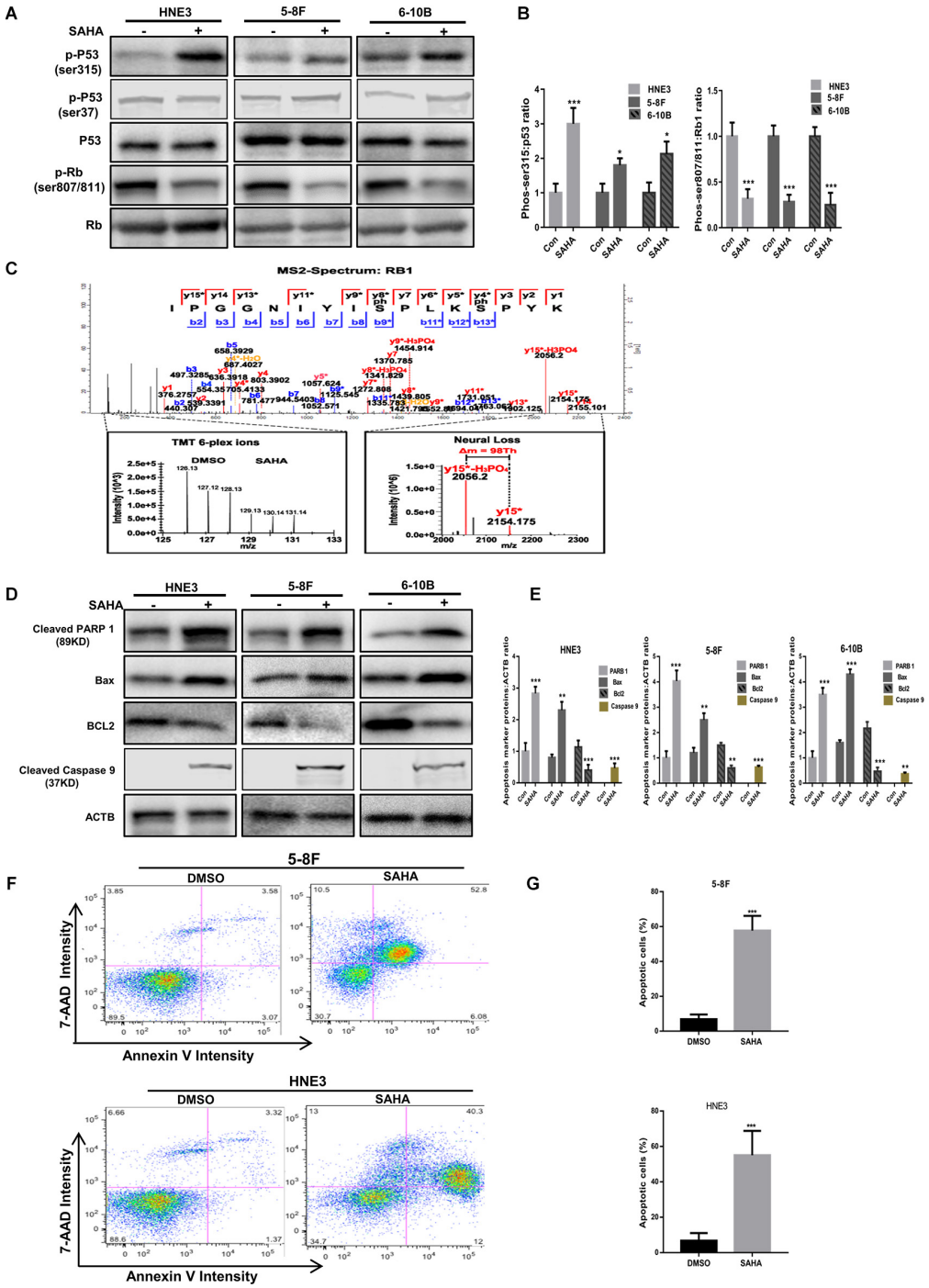


FIGURE 7 | Validation of the activation of p53 pathway and apoptosis by suberoylanilide hydroxamic acid (SAHA). **(A**,**B)** Immunoblot analysis of p53 and Rb1 phosphorylation in nasopharyngeal carcinoma (NPC) cells with or without SAHA treatment. The whole-cell lysates extracted from HNE3, 5-8F, and 6-10B were subjected to Western blot analysis with the indicated antibodies. The protein levels of p53 and Rb1 are shown in **(A)**, while the statistical analysis of all samples is shown in **(B)**. The single asterisk denotes $P < 0.05$. The triple asterisks denote $P < 0.001$. Error bars represent ± SD of triplicate experiments. **(C)** Mapping and quantification of Rb1 phosphopeptide. Representative MS2 spectrum corresponding to a tryptic peptide derived from Rb1 containing two serine phosphorylation sites is presented. **(D,E)** Immunoblot analysis of cleaved PARP 1, cleaved caspase 9, Bax, and Bcl2 in NPC cells with or without SAHA treatment. The whole-cell lysates extracted from HNE3, 5-8F, and 6-10B were subjected to Western blot analysis with the indicated antibodies. The protein levels of cleaved PARP 1, cleaved caspase 9, Bax, and Bcl2 are shown in **(D)**, while the statistical analysis of all samples is shown in **(E)**. The single asterisk denotes $P < 0.05$; the double asterisks denote $P < 0.01$. The triple asterisks denote $P < 0.001$. Error bars represent ± SD of triplicate experiments. **(F,G)** Apoptosis was measured by annexin-V staining in NPC cells. HNE3 and 5-8F were treated with dimethylsulfoxide or SAHA for 24 h and stained with annexin-V-APC and 7-AAD **(F)**. The percentage of apoptotic cells is shown in **(G)**. The triple asterisks denote $P < 0.001$. Error bars represent ± SD of triplicate experiments.

phosphor-peptide enrichment methodology to evaluate protein phosphorylation modulated during SAHA treatment in NPC cells. Our results identified 7,430 phosphorylation sites and 1,836 DEPs, including 847 significantly upregulated and 647 significantly downregulated phosphorylation sites, in response to SAHA stimulation. Several cancer-associated BPs and signaling pathways were enriched, including p53 signaling pathway, cell cycle, DNA damage response, cell metabolism, and cell proliferation *via* GO and IPA pathway analysis. It suggested that SAHA may modulate these BPs and signaling pathways to exert its anti-tumor effect in NPC cells. In agreement with our results, other groups demonstrated that HDAC inhibitors established an anti-tumor activity in head and neck cancer. Francesca et al. revealed that SAHA modulated EGFR receptor expression and reversed epithelial–mesenchymal transition in SCCHN cell lines to display synergistic anti-tumor effects in combination with gefitinib (Bruzzese et al., 2011). It also has reported that TSA suppressed cell migration and invasion through the downregulation of EGFR-Arf1 signaling in SCCHN (He et al., 2019). Moreover, some HDAC inhibitors are in advanced clinical trials either as alone or in combination with other conventional treatments (Han et al., 2015; Caponigro et al., 2016).

Among the upregulated pathways found in this study, p53 signaling pathway ranked top upon SAHA treatment. A comprehensive pathway analysis using IPA database uncovered that p53 phosphorylation at Ser315 residue was elevated upon SAHA treatment, as validated by western blotting. PTMs, including phosphorylation, acetylation, methylation, and sumoylation, regulate the stability and the activity of p53 in response to many stress signals (Luo et al., 2000; Barlev et al., 2001; Buschmann et al., 2001). Phosphorylation at different sites of p53 can have an inhibitory or a stimulatory role in modulating p53-dependent transcription (Bond et al., 1999; Unger et al., 1999; Hirao et al., 2000). Structurally, Ser315 residue localizes in the NLS sequence in the C-terminal region of p53, which contains a tetramerization domain and a regulatory domain (Hupp et al., 1995). Mechanistically, several studies indicated that phosphorylation at Ser315 stimulates its DNA binding function. Jeremy et al. demonstrated that phosphorylation at the Ser315 residue of p53 enhanced p53-dependent transcription through the cdk/cdc-dependent pathway (Blaydes et al., 2001). Additionally, James et al. showed that phosphorylation of p53 at Ser315 site significantly improved the sequence-specific DNA binding activity of p53 protein *in vitro*, possibly by cooperating with other modifications within the C-terminal negative regulatory domain or by alternating N-terminal modifications (Bischoff et al., 1990). Consistent with these findings, our study demonstrated that SAHA boosted the phosphorylation of Ser315 of p53 in NPC cells. However, the upstream signal governing the phosphorylation of Ser315 that is induced by SAHA remains elusive, calling for further investigation.

Due to the crucial suppressive role of p53 protein in the tumorigenesis, we further analyzed the downstream effectors in p53 signaling pathway and found that the phosphorylation of Rb1 at Ser807/811 residues was reduced by SAHA treatment. The retinoblastoma susceptibility gene (Rb1), as the first tumor repressor gene discovered, was described over 30 years ago to have a critical effect on many cellular processes, including cell proliferation, apoptosis, cell cycle, cell differentiation, and DNA repair (Friend et al., 1986; Dowdy et al., 1993; Classon and Harlow, 2002; Dyson, 2016). Phosphorylation and dephosphorylation are well-established mechanisms to regulate Rb1 functions. Dephosphorylated Rb1, an active state, binds and blocks E2F-mediated transcriptional activation. Conversely, Rb1 phosphorylation inactivates its repressive function (Krek et al., 1993; Burke et al., 2010). Ser807/811 locates in the C-terminal of Rb1,with its phosphorylation affecting multiple functions of Rb1, such as Rb1-E2Fs interaction, apoptosis, differentiation, chromosomal stability, and senescence (Ho and Dowdy, 2002; Burkhart and Sage, 2008). Kinases including cdk2/cdk3/cdk5/cdk9 and p38 have been identified to be responsible for Rb1 phosphorylation at Ser807/811 residues. In addition, phosphorylation at Ser807/811 induces structural changes of Rb1, thus impacting phosphorylation at other sites (Brown et al., 1999; Kõivomägi et al., 2011). In addition to these two sites, phosphorylation at many other sites has also been detected according to our quantitative phosphoproteomic data. To obtain a deeper insight on the anti-tumor activity of SAHA in NPC cells, phosphorylation at other sites of Rb1 and the functional outcomes will be investigated in the future.

Nasopharyngeal carcinoma is sensitive to radiotherapy, but the significant rate of relapse and distant metastasis after therapy is still a major cause for NPC lethality (Chang and Adami, 2006). Deep insights and novel therapies are desperately demanded to improve NPC patient prognosis. Chemotherapy combined with immunotherapy or chemotherapy alone demonstrated promising efficacy in several clinical trials for advanced or recurrent NPC (Chen et al., 2018; Ma et al., 2018). Our findings showed that SAHA alone or in combination with current strategies, including radiotherapy, chemotherapy, and immunotherapy, may serve as a potent candidate to combat NPC.

CONCLUSION

In summary, we demonstrated that SAHA inhibits cell proliferation and promotes apoptosis in NPC cells. Taking advantages of TMT labeling, TiO_2 enrichment, and high-resolution LC–MS/MS, we presented the comprehensive quantitative phosphoproteomic profile in NPC cells in response to SAHA treatment, providing a precious resource for further studies on SAHA and NPC. Our results represent a deeper understanding of the underlying mechanism of SAHA's anti-tumor roles. Furthermore, our study indicated that SAHA may serve as a novel therapy for NPC patients.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository with the dataset identifier PXD021029 (http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD021029).

AUTHOR CONTRIBUTIONS

HH, YFu, MYL, and YC designed the study and conceived the experiments. HH and MYL performed the experiments. YZ, FP, YFe, and YFu provided technical and material support. HH, MLL, LC, and ZC analyzed and interpreted the data. HH and MYL wrote the manuscript. All the authors read and approved the final manuscript.

FUNDING

The present study was supported by the National Natural Science Foundation of China (Nos. 81974074, 81570537, and 81902858) and the Major State Basic Research Development Program of China (No. 2014CBA02000-4).

ACKNOWLEDGMENTS

We thank Dr. Jun Zhao for the constructive comments on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.577784/full#supplementary-material

REFERENCES

Adham, M., Kurniawan, A. N., Muhtadi, A. I., Roezin, A., Hermani, B., Gondhowiardjo, S., et al. (2012). Nasopharyngeal carcinoma in Indonesia: epidemiology, incidence, signs, and symptoms at presentation. *Chin. J. Cancer* 31, 185–196. doi: 10.5732/cjc.011.10328

Barlev, N. A., Liu, L., Chehab, N. H., Mansfield, K., Harris, K. G., Halazonetis, T. D., et al. (2001). Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol. Cell* 8, 1243–1254. doi: 10.1016/s1097-2765(01)00414-2

Bischoff, J. R., Friedman, P. N., Marshak, D. R., Prives, C., and Beach, D. (1990). Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4766–4770. doi: 10.1073/pnas.87.12.4766

Blaydes, J. P., Luciani, M. G., Pospisilova, S., Ball, H. M., Vojtesek, B., and Hupp, T. R. (2001). *Stoichiometric phosphorylation* of human p53 at Ser315 stimulates p53-dependent transcription. *J. Biol. Chem.* 276, 4699–4708. doi: 10.1074/jbc.M003485200

Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., et al. (2004). Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 101, 17033–17038. doi: 10.1073/pnas.0407902101

Bond, J. A., Webley, K., Wyllie, F. S., Jones, C. J., Craig, A., Hupp, T., et al. (1999). p53-Dependent growth arrest and altered p53-immunoreactivity following metabolic labelling with 32P ortho-phosphate in human fibroblasts. *Oncogene* 18, 3788–3792. doi: 10.1038/sj.onc.1202733

Brown, V. D., Phillips, R. A., and Gallie, B. L. (1999). Cumulative effect of phosphorylation of pRB on regulation of E2F activity. *Mol. Cell Biol.* 19, 3246–3256. doi: 10.1128/mcb.19.5.3246

Bruzzese, F., Leone, A., Rocco, M., Carbone, C., Piro, G., Caraglia, M., et al. (2011). HDAC inhibitor vorinostat enhances the antitumor effect of gefitinib in squamous cell carcinoma of head and neck by modulating ErbB receptor expression and reverting EMT. *J. Cell Physiol.* 226, 2378–2390. doi: 10.1002/jcp.22574

Burke, J. R., Deshong, A. J., Pelton, J. G., and Rubin, S. M. (2010). Phosphorylation-induced conformational changes in the retinoblastoma protein inhibit E2F transactivation domain binding. *J. Biol. Chem.* 285, 16286–16293. doi: 10.1074/jbc.M110.108167

Burkhart, D. L., and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* 8, 671–682. doi: 10.1038/nrc2399

Buschmann, T., Potapova, O., Bar-Shira, A., Ivanov, V. N., Fuchs, S. Y., Henderson, S., et al. (2001). Jun NH2-terminal kinase phosphorylation of p53 on Thr-81 is important for p53 stabilization and transcriptional activities in response to stress. *Mol. Cell Biol.* 21, 2743–2754. doi: 10.1128/mcb.21.8.2743-2754.2001

Cao, S. M., Simons, M. J., and Qian, C. N. (2011). The prevalence and prevention of nasopharyngeal carcinoma in China. *Chin. J. Cancer* 30, 114–119. doi: 10.5732/cjc.010.10377

Caponigro, F., Di Gennaro, E., Ionna, F., Longo, F., Aversa, C., Pavone, E., et al. (2016). Phase II clinical study of valproic acid plus cisplatin and cetuximab in recurrent and/or metastatic squamous cell carcinoma of head and neck-V-CHANCE trial. *BMC Cancer* 16:918. doi: 10.1186/s12885-016-2957-y

Chang, E. T., and Adami, H. O. (2006). The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiol. Biomark. Prev.* 15, 1765–1777. doi: 10.1158/1055-9965.epi-06-0353

Chen, Y. P., Tang, L. L., Yang, Q., Poh, S. S., Hui, E. P., Chan, A. T. C., et al. (2018). Induction chemotherapy plus concurrent chemoradiotherapy in endemic nasopharyngeal carcinoma: individual patient data pooled analysis of four randomized trials. *Clin. Cancer Res.* 24, 1824–1833. doi: 10.1158/1078-0432.ccr-17-2656

Cicenas, J. (2008). The potential role of Akt phosphorylation in human cancers. *Int. J. Biol. Markers* 23, 1–9. doi: 10.5301/jbm.2008.618

Classon, M., and Harlow, E. (2002). The retinoblastoma tumour suppressor in development and cancer. *Nat. Rev. Cancer* 2, 910–917. doi: 10.1038/nrc950

Dowdy, S. F., Hinds, P. W., Louie, K., Reed, S. I., Arnold, A., and Weinberg, R. A. (1993). Physical interaction of the retinoblastoma protein with human D cyclins. *Cell* 73, 499–511. doi: 10.1016/0092-8674(93)90137-f

Du, C., Zhang, C., Hassan, S., Biswas, M. H., and Balaji, K. C. (2010). Protein kinase D1 suppresses epithelial-to-mesenchymal transition through phosphorylation of snail. *Cancer Res.* 70, 7810–7819. doi: 10.1158/0008-5472.can-09-4481

Duvic, M., Talpur, R., Ni, X., Zhang, C., Hazarika, P., Kelly, C., et al. (2007). Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL). *Blood* 109, 31–39. doi: 10.1182/blood-2006-06-025999

Dyson, N. J. (2016). RB1: a prototype tumor suppressor and an enigma. *Genes Dev.* 30, 1492–1502. doi: 10.1101/gad.282145.116

Friend, S. H., Bernards, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M., et al. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323, 643–646. doi: 10.1038/323643a0

Gubern, A., Joaquin, M., Marques, M., Maseres, P., Garcia-Garcia, J., Amat, R., et al. (2016). The N-terminal phosphorylation of RB by p38 bypasses its inactivation by CDKs and prevents proliferation in cancer cells. *Mol. Cell* 64, 25–36. doi: 10.1016/j.molcel.2016.08.015

Han, J. Y., Lee, S. H., Lee, G. K., Yun, T., Lee, Y. J., Hwang, K. H., et al. (2015). Phase I/II study of gefitinib (Iressa(®)) and vorinostat (IVORI) in previously treated patients with advanced non-small cell lung cancer. *Cancer Chemother. Pharmacol.* 75, 475–483. doi: 10.1007/s00280-014-2664-9

He, L., Gao, L., Shay, C., Lang, L., Lv, F., and Teng, Y. (2019). Histone deacetylase inhibitors suppress aggressiveness of head and neck squamous cell carcinoma via histone acetylation-independent blockade of the EGFR-Arf1 axis. *J. Exp. Clin. Cancer Res.* 38:84. doi: 10.1186/s13046-019-1080-8

Hirao, A., Kong, Y. Y., Matsuoka, S., Wakeham, A., Ruland, J., Yoshida, H., et al. (2000). DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 287, 1824–1827. doi: 10.1126/science.287.5459.1824

Ho, A., and Dowdy, S. F. (2002). Regulation of G(1) cell-cycle progression by oncogenes and tumor suppressor genes. *Curr. Opin. Genet. Dev.* 12, 47–52. doi: 10.1016/s0959-437x(01)00263-5

Hoffman, M. D., Sniatynski, M. J., and Kast, J. (2008). Current approaches for global post-translational modification discovery and mass spectrometric analysis. *Anal. Chim. Acta* 627, 50–61. doi: 10.1016/j.aca.2008.03.032

Hupp, T. R., Sparks, A., and Lane, D. P. (1995). Small peptides activate the latent sequence-specific DNA binding function of p53. *Cell* 83, 237–245. doi: 10.1016/0092-8674(95)90165-5

Kõivomägi, M., Valk, E., Venta, R., Iofik, A., Lepiku, M., Balog, E. R., et al. (2011). Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase. *Nature* 480, 128–131. doi: 10.1038/nature10560

Komatsu, N., Kawamata, N., Takeuchi, S., Yin, D., Chien, W., Miller, C. W., et al. (2006). SAHA, a HDAC inhibitor, has profound anti-growth activity against non-small cell lung cancer cells. *Oncol. Rep.* 15, 187–191.

Konstantinopoulos, P. A., Wilson, A. J., Saskowski, J., Wass, E., and Khabele, D. (2014). Suberoylanilide hydroxamic acid (SAHA) enhances olaparib activity by targeting homologous recombination DNA repair in ovarian cancer. *Gynecol. Oncol.* 133, 599–606. doi: 10.1016/j.ygyno.2014.03.007

Krek, W., Livingston, D. M., and Shirodkar, S. (1993). Binding to DNA and the retinoblastoma gene product promoted by complex formation of different E2F family members. *Science* 262, 1557–1560. doi: 10.1126/science.8248803

Krueger, K. E., and Srivastava, S. (2006). Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics. *Mol. Cell Proteom.* 5, 1799–1810. doi: 10.1074/mcp.R600009-MCP200

Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jørgensen, T. J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell Proteom.* 4, 873–886. doi: 10.1074/mcp.T500007-MCP200

Li, D. J., Deng, G., Xiao, Z. Q., Yao, H. X., Li, C., Peng, F., et al. (2009). Identificating 14-3-3 sigma as a lymph node metastasis-related protein in human lung squamous carcinoma. *Cancer Lett.* 279, 65–73. doi: 10.1016/j.canlet.2009.01.028

Luo, J., Su, F., Chen, D., Shiloh, A., and Gu, W. (2000). Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature* 408, 377–381. doi: 10.1038/35042612

Ma, B. B. Y., Lim, W. T., Goh, B. C., Hui, E. P., Lo, K. W., Pettinger, A., et al. (2018). Antitumor activity of nivolumab in recurrent and metastatic nasopharyngeal carcinoma: an international, multicenter study of the mayo clinic phase 2 consortium (NCI-9742). *J. Clin. Oncol.* 36, 1412–1418. doi: 10.1200/jco.2017.77.0388

Marks, P. A., and Breslow, R. (2007). Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* 25, 84–90. doi: 10.1038/nbt1272

Minucci, S., and Pelicci, P. G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat. Rev. Cancer* 6, 38–51. doi: 10.1038/nrc1779

Munster, P. N., Troso-Sandoval, T., Rosen, N., Rifkind, R., Marks, P. A., and Richon, V. M. (2001). The histone deacetylase inhibitor suberoylanilide hydroxamic acid induces differentiation of human breast cancer cells. *Cancer Res.* 61, 8492–8497.

Olsen, E. A., Kim, Y. H., Kuzel, T. M., Pacheco, T. R., Foss, F. M., Parker, S., et al. (2007). Phase IIb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. *J. Clin. Oncol.* 25, 3109–3115. doi: 10.1200/jco.2006.10.2434

Roberts, P. J., and Der, C. J. (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* 26, 3291–3310. doi: 10.1038/sj.onc.1210422

Sanidas, I., Morris, R., Fella, K. A., Rumde, P. H., Boukhali, M., Tai, E. C., et al. (2019). A code of mono-phosphorylation modulates the function of RB. *Mol. Cell* 73, 985–1000.e1006. doi: 10.1016/j.molcel.2019.01.004

Shahbazian, M. D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* 76, 75–100. doi: 10.1146/annurev.biochem.76.052705.162114

Shawver, L. K., Slamon, D., and Ullrich, A. (2002). Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer Cell* 1, 117–123. doi: 10.1016/s1535-6108(02)00039-9

Sheehan, K. M., Calvert, V. S., Kay, E. W., Lu, Y., Fishman, D., Espina, V., et al. (2005). Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol. Cell Proteom.* 4, 346–355. doi: 10.1074/mcp.T500003-MCP200

Siegel, D., Hussein, M., Belani, C., Robert, F., Galanis, E., Richon, V. M., et al. (2009). Vorinostat in solid and hematologic malignancies. *J. Hematol. Oncol.* 2:31. doi: 10.1186/1756-8722-2-31

Unger, T., Sionov, R. V., Moallem, E., Yee, C. L., Howley, P. M., Oren, M., et al. (1999). Mutations in serines 15 and 20 of human p53 impair its apoptotic activity. *Oncogene* 18, 3205–3212. doi: 10.1038/sj.onc.1202656

Wagner, J. M., Hackanson, B., Lübbert, M., and Jung, M. (2010). Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin. Epigenet.* 1, 117–136. doi: 10.1007/s13148-010-0012-4

Wu, Q., Cheng, Z., Zhu, J., Xu, W., Peng, X., Chen, C., et al. (2015). Suberoylanilide hydroxamic acid treatment reveals crosstalks among proteome, ubiquitylome and acetylome in non-small cell lung cancer A549 cell line. *Sci. Rep.* 5:9520. doi: 10.1038/srep09520

Xu, G., Wang, J., Wu, Z., Qian, L., Dai, L., Wan, X., et al. (2014). SAHA regulates histone acetylation, Butyrylation, and protein expression in neuroblastoma. *J. Proteome Res.* 13, 4211–4219. doi: 10.1021/pr500497e

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 © 2020 Huang, Fu, Zhang, Peng, Lu, Feng, Chen, Chen, Li and Chen. 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 December 2020
doi: 10.3389/fcell.2020.545934

Proteome-Wide Alterations of Asymmetric Arginine Dimethylation Associated With Pancreatic Ductal Adenocarcinoma Pathogenesis

Meijin Wei[1†], Chaochao Tan[2,3†], Zhouqin Tang[1], Yingying Lian[1], Ying Huang[4], Yi Chen[1], Congwei Chen[1], Wen Zhou[1], Tao Cai[5] and Jiliang Hu[1]**

[1] School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China, [2] Department of Clinical Laboratory, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, China, [3] Translational Medicine Research Institute, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, China, [4] Department of Emergency, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, China, [5] Department of Neurosurgery, The Third Xiangya Hospital, Central South University, Changsha, China

Arginine methylation catalyzed by protein arginine methyltransferases (PRMTs) performs essential roles in regulating cancer initiation and progression, but its implication in pancreatic ductal adenocarcinoma (PDAC) requires further elucidation. In this study, asymmetric dimethylarginine (ADMA)-containing peptides in PDAC cell line PANC-1 were identified by label-free quantitative proteomics combined with affinity purification, using human non-cancerous pancreatic ductal epithelium cell line HPDE6c7 as the control. In total, 289 ADMA sites in 201 proteins were identified in HPDE6c7 and PANC-1 cells, including 82 sites with lower dimethylation and 37 sites with higher dimethylation in PANC-1 cells compared with HPDE6c7 cells. These ADMA-containing peptides demonstrated significant enrichment of glycine and proline residues in both cell lines. Importantly, leucine residues were significantly enriched in ADMA-containing peptides identified only in HPDE6c7 cells or showing lower dimethylation in PANC-1 cells. ADMA-containing proteins were significantly enriched in multiple biological processes and signaling cascades associated with cancer development, such as spliceosome machinery, the Wnt/β-catenin, Hedgehog, tumor growth factor beta (TGF-β), and mitogen-activated protein kinase (MAPK) signaling pathways. Moreover, PDAC cell lines with enhanced cell viability showed lower PRMT4 protein abundance and global ADMA-containing protein levels compared with HPDE6c7. PRMT4 overexpression partially recovered ADMA-containing protein levels and repressed viability in PANC-1 cells. These results revealed significantly altered ADMA-containing protein profiles in human pancreatic carcinoma cells, which provided a basis for elucidating the pathogenic roles of PRMT-mediated protein methylation in pancreatic cancer.

Keywords: protein methylation, asymmetric dimethylarginine, pancreatic ductal adenocarcinoma, PRMT4, label-free quantitative proteomics

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Rainer Böger, University Medical Center Hamburg-Eppendorf, Germany
Adam Frankel, University of British Columbia, Canada

****Correspondence:***
Tao Cai
1172070101@qq.com
Jiliang Hu
jlhu@gzucm.edu.cn
orcid.org/0000-0001-6754-3788

† These authors have contributed equally to this work

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

Received: *26 March 2020*
Accepted: *06 November 2020*
Published: *03 December 2020*

Citation:
Wei M, Tan C, Tang Z, Lian Y, Huang Y, Chen Y, Chen C, Zhou W, Cai T and Hu J (2020) Proteome-Wide Alterations of Asymmetric Arginine Dimethylation Associated With Pancreatic Ductal Adenocarcinoma Pathogenesis. Front. Cell Dev. Biol. 8:545934. doi: 10.3389/fcell.2020.545934

INTRODUCTION

Pancreatic cancer is a common malignant disorder with rapid progression and poor prognosis and remains one of the leading causes of cancer related deaths worldwide (Chen et al., 2016; Siegel et al., 2018). Pancreatic ductal adenocarcinoma (PDAC) is a major pancreatic cancer subtype, accounting for more than 85% of global pancreatic cancer cases, with a 5-year survival rate of less than 5% (Ryan et al., 2014; Siegel et al., 2018). PDAC pathogenesis is driven by multiple genetic alterations such as the activating mutation of KRAS (v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog) (Ryan et al., 2014; Buscail et al., 2020). However, current therapeutic regimens targeting KRAS have failed to lower PDAC mortality and improve prognosis, partially due to limited effectiveness (Iovanna and Dusetti, 2017; Buscail et al., 2020). In recent years, post-translational modifications (PTMs) have emerged as essential regulators of PDAC initiation and progression (Roth et al., 2017; Tan et al., 2019; Pan et al., 2020). The identification and functional investigation of protein modifications associated with PDAC pathogenesis could provide alternative targets for pancreatic cancer diagnosis and treatment.

Arginine methylation refers to the addition of methyl groups onto guanidino groups localized on the side chains of protein arginine (Arg) residues, and was first discovered as a protein modification type in 1971 (Bedford and Clarke, 2009; Blanc and Richard, 2017; Guccione and Richard, 2019; Jarrold and Davies, 2019). Arginine methylation is catalyzed by the protein arginine methyltransferases (PRMTs), resulting in the formation of monomethylarginine (MMA), asymmetrical dimethylarginine (ADMA), or symmetrical dimethylarginine (SDMA) (Guccione and Richard, 2019; Jarrold and Davies, 2019). Methylation increases the hydrophobicity and bulkiness of arginine residues in modified proteins and consequently interferes with their interactions with other proteins or nucleic acid partners (Shishkova et al., 2017). Arginine methylation is closely involved in various biological processes through affecting gene transcription, pre-mRNA splicing, protein translation, and synthesis (Guccione and Richard, 2019; Jarrold and Davies, 2019). Importantly, arginine methylation deregulation is also closely associated with cancer initiation, metastasis, and drug resistance (Yang and Bedford, 2013; Guccione and Richard, 2019; Jarrold and Davies, 2019). For instance, the asymmetrical dimethylation of histone H4 at position R3 (H4R3me2a) by PRMT1 mediates the epigenetic reprogramming and aberrant transcriptional regulation during the progression of acute myeloid leukemia (Cheung et al., 2016). Arginine methylation on histone tails have been established as key epigenetic events that drive cancer development and progression (Waldmann and Schneider, 2013; Jarrold and Davies, 2019).

Recent research has showed that arginine methylation in non-histone proteins also performs pivotal roles in cancer pathogenesis (Biggar and Shawn, 2015). For instance, PRMT4, alternatively known as co-activator-associated arginine methyltransferase 1 (CARM1), is overexpressed in breast cancer cells and regulates breast cancer progression and chemo-sensitivity through catalyzing the arginine methylation of multiple proteins such as pyruvate kinase M2 (PKM2) isoform, BAF155, and RNA polymerase II mediator complex subunit 12 (MED12) (Wang et al., 2014, 2015; Liu et al., 2017). However, the expression of PRMT4 protein is significantly suppressed in pancreatic cancer cells, resulting in reduced asymmetric arginine dimethylation of malate dehydrogenase 1 (MDH1) and enhanced non-canonical glutamine metabolism (Wang Y. P. et al., 2016). Moreover, asymmetric arginine dimethylation of other non-histone proteins such as Gli1 (glioma-associated oncogene homolog) and ATP-binding cassette subfamily G member 2 (ABCG2) were also implicated in pancreatic cancer pathogenesis, suggesting the important roles of asymmetric arginine dimethylation in PDAC (Wang Y. et al., 2016; Hsu et al., 2018). However, current knowledge of asymmetric arginine dimethylation in PDAC pathogenesis remains limited due to the lack of large-scale characterization of ADMA-containing proteins in pancreatic cancer cells (Pan et al., 2020).

Mass spectrometry-based proteomics combined with immunoaffinity purification has been successfully applied for profiling protein arginine methylation in *Trypanosoma brucei* (Fisk et al., 2013), *Plasmodium falciparum* (Zeeshan et al., 2017), *Toxoplasma gondii* (Yakubu et al., 2017), human renal epithelial cells (Sylvestersen et al., 2014), T cells (Geoghegan et al., 2015), breast and colon cancer cells (Guo et al., 2014; Shishkova et al., 2017). In this study, we performed a global characterization of ADMA-containing proteins in human pancreatic ductal epithelial cells and PDAC cells through label-free quantitative proteomics coupled with affinity purification, which laid a foundation for elucidating the roles of PRMT-mediated protein methylation in PDAC pathogenesis.

MATERIALS AND METHODS

Cell Culture, Transfection, and Viability

The immortalized human pancreatic ductal epithelium cell line HPDE6c7 was obtained from the Kyushu University (Japan) and two PDAC cell lines PANC-1 and BxPC-3 were purchased from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China). Short tandem repeat (STR) profiling was used to authenticate cell identity. Cells were cultured in DMEM (Dulbecco Modified Eagle Medium) containing 10% fetal bovine serum (Invitrogen) and penicillin/streptomycin at 37°C with 5% CO_2. Human PRMT4 gene coding sequences containing HA-tag sequence at N-terminal were amplified by RT-PCR and ligated with the Lv-CMV-EGFP vector (Cyagen Biosciences, Suzhou, China), and the packaged recombinant lentivirus vectors were transfected into PANC-1 cells for stable expression of HA-PRMT4 proteins, as previously described (Wu et al., 2012). PANC-1 cells transfected with packaged Lv-CMV-EGFP vectors without ligation of PRMT4 gene coding sequences were used as the control group. Cell viabilities were measured by the CCK-8 (Cell Counting Kit-8; Dojindo, Japan) for three biological replicates as previously introduced (Tan et al., 2019).

Protein Extraction and Quality Control

Total proteins were extracted from approximately 2×10^8 cells at 80% confluency using a lysis buffer containing 100 mM

HEPES (pH 8.0), 8 M urea, and 1% protease inhibitor cocktail (Sigma-Aldrich). After sonication and centrifuged at 20,000 × *g* for 15 min, the concentrations of proteins in the supernatant fraction were determined by the bicinchoninic acid (BCA) method. Protein quality evaluation was performed by Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) separation combined with Coomassie brilliant blue staining, and the numbers of protein groups were identified by a preliminary liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

Tryptic Digestion

Total proteins (15 mg/group) were mixed with 1.25 M DTT (dithiothreitol; final concentration:10 mM) and incubated at 37°C for 30 min with gentle shaking, followed by incubation with 50 mM iodoacetamide (IAM) in darkness for 30 min at room temperature. Subsequently, proteins were digested in a solution with trypsin (Sigma-Aldrich) at 37°C for 17 h, as previously described (Hu et al., 2015). The resulting peptide solution was mixed with 0.1% trifluoroacetic acid (TFA), desalted by solid-phase extraction on a C18 cartridge column, and lyophilized for approximately 2 days to remove TFA, as previously described (Hu et al., 2015).

ADMA-Containing Peptide Enrichment

The enrichment of ADMA-containing peptides was done using the PTMScan® Asymmetric Di-Methyl Arginine Motif Kit (#13474; Cell Signaling Technology), following the manufacturer's instructions. Briefly, peptide powders were resuspended with 1.4 mL pre-chilled IAP buffer and incubated with one vial of immunoaffinity beads for 1.5 h at 4°C. After being centrifuged at 2,000 × *g* for 30 s, the supernatant was removed, and immunoaffinity beads were washed three times with pre-chilled IAP buffer and washed three times with pre-chilled ddH_2O. Subsequently, immunoaffinity beads were eluted twice with 40 μL 0.15% TFA by incubation at room temperature for 10 min. The resulting peptide solution was desalted on a C18-StageTips column and dried by lyophilization in a vacuum concentrator for mass spectrometry.

LC-MS/MS

Asymmetric dimethylarginine-containing peptide powders were then dissolved in 5% acetonitrile containing 0.1% TFA, which were subjected to LC-MS/MS (liquid chromatography coupled with tandem mass spectrometry) using a Thermo Scientific Q Exactive MS system coupled online to an Easy-nLC 1000 instrument. Peptide solution was first loaded into the Thermo Scientific Acclaim PepMap100 loading column (100 μm × 20 mm, nanoViper C18) through the autosampler, followed by separation on a Thermo Scientific EASY column (75 μm × 250 mm) packed with C18-A2 particles (3 μm). Buffer A (0.1% FA) and Buffer B (84% acetonitrile and 0.1% FA) were used as the mobile phases. The flow rate was set to 300 nL/min using the following non-linear gradient: 5–8% buffer B, 10 min; 8–20% buffer B, 1 h; 20–30% buffer B, 7 min; 30–100% buffer B, 3 min; 100% buffer B, 12 min. Tandem mass spectrometry was performed in the positive ion mode, and precursor ions ranging from 300 to 1800 m/z were measured at 70,000 resolution (200 m/z) with an automatic gain control target of 10^6 ions and a maximum injection time of 10 ms. Twenty precursor ions with the highest intensities from each full scan were selected for fragmentation by higher-energy C-trap dissociation (HCD) (normalized collision energy: 30 eV). MS2 spectra were acquired at a resolution of 17,500 at 200 m/z, and the isolation window was set to 2 m/z. All mass spectrometry data were deposited to the ProteomeXchange Consortium using the iProX partner repository (Ma et al., 2019) with the dataset identifier PXD017577[1].

Database Searching

Raw data from LC/MS/MS analysis were searched against the Uniprot proteome database released on 30/1/2018 (taxonomy: Homo sapiens) containing 20,244 canonical and isoform entries using the MaxQuant software (version: 1.6.0.16)[2]. The protease for protein digestion was set to trypsin and the maximum allowable mis-cleavages were set to two. A minimum peptide length of 7 amino acids and a maximum peptide mass of 4,600 Da were used for the database search. The mass tolerances for both precursor ions and fragment ions were 20 PPM. Cysteine carbamidomethylation was used as the fixed modification and variable modifications included N-terminal acetylation, methionine oxidation, and arginine dimethylation (28.0313 Da). A false discovery rate (FDR) of <0.01 was used for both peptide and protein identification. An Andromeda score of >40 and a localization probability of >0.75 were used for acceptance of peptide modification. Other parameters were set to default values.

Quantitation and Bioinformatics

The differential levels of ADMA-containing peptides between the HPDE6c7 and PANC-1 cells were analyzed by Perseus software (version 1.6.6.0)[3] following the manufacturer's instructions (Tyanova et al., 2016). Briefly, the "Dimethyl (R)Sites.txt" file produced from database search was loaded to the Perseus platform, followed by filtering out contaminants and reverse (decoy) database hits. ADMA sites with a localization probability of <0.75 were excluded. The peptide intensities were logarithm-transformed using the formula "log2(x)" followed by median normalization. The Pearson correlation between biological replicates were >0.977 in the HPDE6c7 group and >0.785 in the PANC-1 group. After adding the missing data, we defined the significantly differential ADMA-containing peptides by the Student's *t*-test ($P < 0.05$), which were then used for hierarchical clustering and PCA analysis. Consensus sequences of ADMA-containing peptides were predicted by the online pLogo platform[4] as previously introduced (O'Shea et al., 2013). We analyzed functional categorization based on Gene ontology (GO) and Kyoto Encyclopedia of Genes

[1]http://proteomecentral.proteomexchange.org

[2]https://www.maxquant.org/

[3]https://www.maxquant.org/perseus/

[4]https://plogo.uconn.edu/

and Genomes (KEGG) signaling pathways using the DAVID database (Huang da et al., 2009). The signaling pathway diagrams of ADMA-containing proteins were modified from the KEGG database[5].

Immunoblotting

Protein abundances and total ADMA-containing protein levels between HPDE6c7 and PANC-1 cells were detected by western blotting, as previously introduced (Hu et al., 2017). Primary antibodies targeting PRMT4 (#4438; Cell Signaling Technology), ADMA (#13522; Cell Signaling Technology), and β-actin (#ab8227; Abcam) were used in this study. At least three biological replicates were done for protein quantitation.

Experimental Design and Statistical Rationale

HPDE6c7 and PANC-1 cells cultured under the same conditions were subjected to ADMA-containing peptide identification by label-free quantitative proteomics. Three biological replicates were performed in each cell line ($n = 3$). ADMA-containing peptides enriched from each sample were continuously analyzed by LC-MS/MS. All raw data were searched together against the Uniprot proteome database using the MaxQuant software. The FDR of <0.01 and Andromeda score of >40 were used as thresholds for ADMA-containing peptide and protein identification. Differentially methylated peptides were defined by a P-value of <0.05 from Student's t-test using the Perseus software. Significant differences in other assays ($P < 0.05$) were evaluated by the Student's t-test or analysis of variance (ANOVA) using the SPSS 20.0 software.

RESULTS

Global Profiling of ADMA-Containing Proteins in HPDE6c7 and PANCI-1 Cells

For the profiling of asymmetric arginine dimethylation, ADMA-containing peptides in both HPDE6c7 and PANC-1 cells were identified by label-free quantitative proteomics following immunoaffinity purification (**Figure 1A**). In total, 289 ADMA sites were identified by three biological replicates, including 243 and 239 ADMA sites, which were identified in the HPDE6c7 and PANC-1 cells, respectively (**Figures 1B,C** and **Supplementary Table 1**). Among them, 193 ADMA sites were identified in both cell lines, while 50 and 46 ADMA sites were identified exclusively in HPDE6c7 and PDAC cells, respectively (**Figure 1D**). These ADMA-containing peptides were also mapped to a total of 201 methylated proteins, including 174 and 164 methylated proteins in HPDE6c7 and PDAC cells, respectively (**Figure 1E** and **Supplementary Table 1**). Among them, the asymmetric dimethylation of 50.75% (102/201) proteins were previously reported, validating the ADMA-containing protein dataset (**Figure 1F** and **Supplementary Table 1**). The previously reported asymmetric dimethylation of TATA-binding protein associated factor 15 (TAF15) protein at Arg206 (Jobert et al., 2009), was also detected in our proteomic analysis (**Figure 1G**). Moreover, 99 ADMA-containing proteins were newly identified in this study, and their possible roles in PDAC pathogenesis deserve further investigation in future.

Differential Asymmetric Arginine Dimethylation in Pancreatic Cancer Cells

Label-free quantitation of ADMA-containing peptides was subsequently performed using the Perseus software. We found that 119 of the total 193 ADMA sites identified in both the HPDE6c7 and PANC-1 cells, were differentially dimethylated between these two cell lines (**Figure 2A** and **Supplementary Table 2**). Among them, 82 ADMA sites showed lower dimethylation levels in PANC-1 cells compared with HPDE6c7 cells, which is much more than those showing higher dimethylation in PANC-1 cells (**Figure 2A**). The dimethylation of 58 proteins in PANC-1 cells was significantly lower than the HPDE6c7 cells, while only 28 proteins showed higher dimethylation in PANC-1 cells (**Figure 2B**). Our hierarchical clustering also demonstrated significantly differential arginine dimethylation between HPDE6c7 and PANC-1 cells, and the majority of differential ADMA-containing proteins showed lowered dimethylation levels in PANC-1 cells compared with HPDE6c7 cells (**Figure 2C**). Moreover, biological replicates of HPDE6c7 and PANC-1 cells were separated in the principal component analysis (PCA), based on differentially dimethylated proteins (**Figure 2D**). Together, these results show significantly differential ADMA-containing protein profiles between the HPDE6c7 and PANC-1 cells.

Lower Leucine Enrichment Flanking ADMA Sites in PDAC Cells

Previous reports have revealed significant enrichment of proline and glycine residues in the vicinity of ADMA sites (Uhlmann et al., 2012; Fisk et al., 2013; Guo et al., 2014; Shishkova et al., 2017). With this large set of ADMA-containing peptides, we also characterized consensus sequences of asymmetric arginine dimethylation using the pLogo software. We showed that proline (P) at positions −1 and glycine at multiple positions were significantly overrepresented in ADMA-containing peptides in both HPDE6c7 and PANC-1 cells (**Figure 3A** and **Supplementary Table 1**). This observation is consistent with previous reports (Uhlmann et al., 2012; Fisk et al., 2013; Guo et al., 2014; Shishkova et al., 2017), suggesting unbiased ADMA-containing peptide identification in our proteomic analysis. More importantly, we showed that leucine residue at position +3 was preferentially present in ADMA-containing peptides, especially in those identified in HPDE6c7 cells and differentially dimethylated between HPDE6c7 and PANC-1 cells (**Figures 3A,B** and **Supplementary Table 1**). The frequency of leucine at position +3 was even higher than proline and glycine in ADMA-containing peptides with lower dimethylation in PANC-1 cells, but

[5]www.kegg.jp

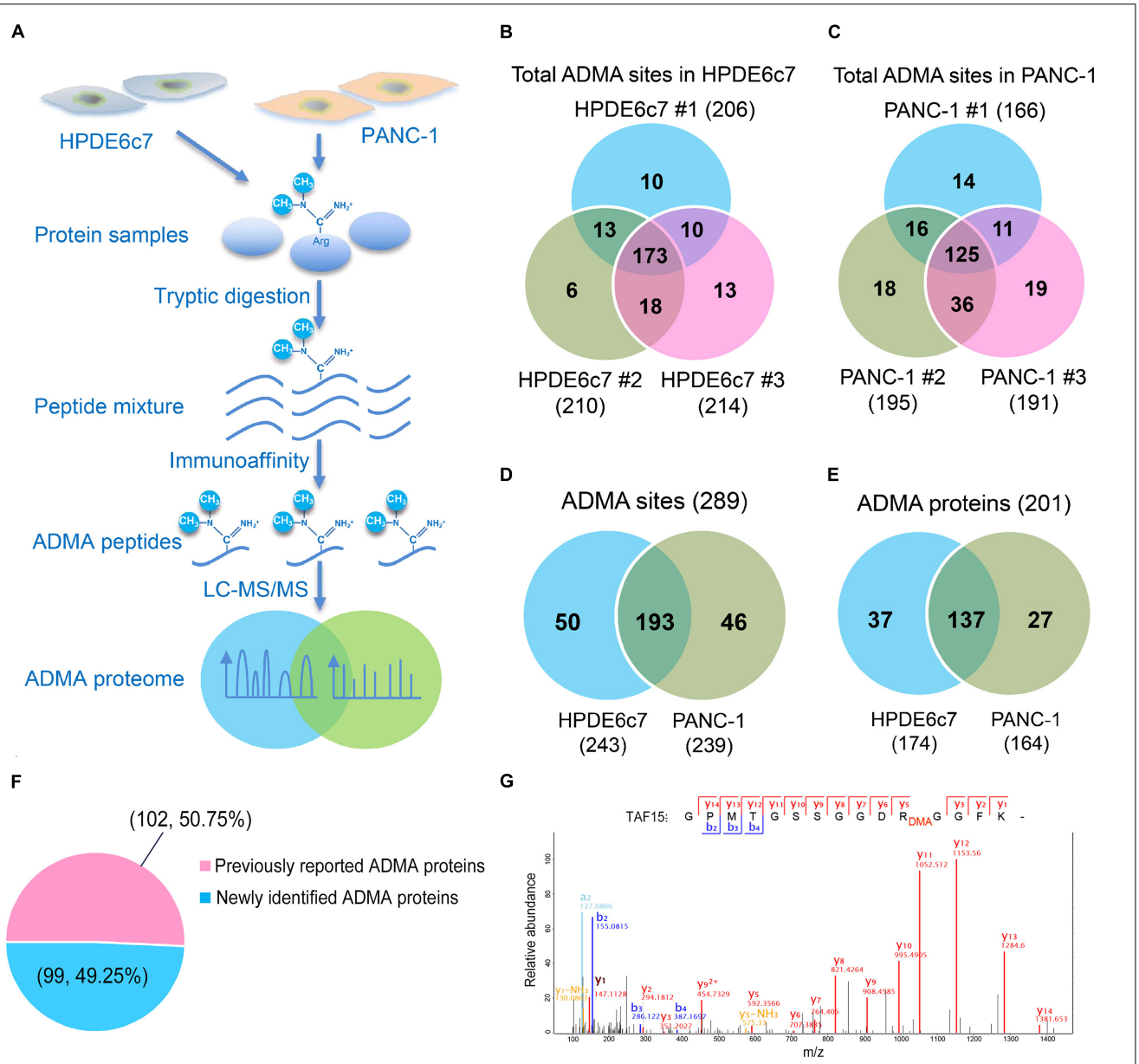


FIGURE 1 | Mass spectrometric identification of ADMA-containing peptides in HPDE6c7 and PDAC cells. **(A)** A schematic illustration of proteomic characterization of ADMA-containing peptides in pancreatic cells. ADMA-containing peptides in PANC-1 cells were enriched by immunoaffinity and identified by LC-MS/MS, and human normal pancreatic ductal epithelium cell line HPDE6c7 was used as the control. **(B,C)** ADMA sites identified in HPDE6c7 and PANC-1 cells by three biological replicates. ADMA site numbers in HPDE6c7 **(B)** and PANC-1 **(C)** cells were shown in Venn diagrams. **(D)** Total numbers of ADMA sites in HPDE6c7 and PANC-1 cells. **(E)** Total numbers of ADMA-containing proteins in HPDE6c7 and PANC-1 cells. **(F)** Percentage of previously reported and newly identified ADMA-containing proteins. The numbers of previously reported or newly identified ADMA-containing proteins were shown in the brackets together with their percentages. **(G)** Mass spectrometric identification of TAF15 protein asymmetric dimethylation at Arg206 in PANC-1 cells.

not in those with higher dimethylation in PANC-1 cells (**Figure 3B** and **Supplementary Table 2**). Consistently, the enrichment of leucine residues was observed in ADMA-containing peptides identified only in HPDE6c7 cells, other than those identified only in PANC-1 cells (**Figure 3C** and **Supplementary Table 1**). These results suggested the differential properties of ADMA-flanking residues between HPDE6c7 and PANC-1 cells.

Functional Categorization of ADMA-Containing Proteins

To explore the potential biological roles of asymmetric arginine dimethylation, we then performed a functional categorization of ADMA-containing proteins in HPDE6c7 and PANC-1 cells based on Gene Ontology (GO) terms. In general, these modified proteins were significantly enriched in several biological processes associated with RNA processing

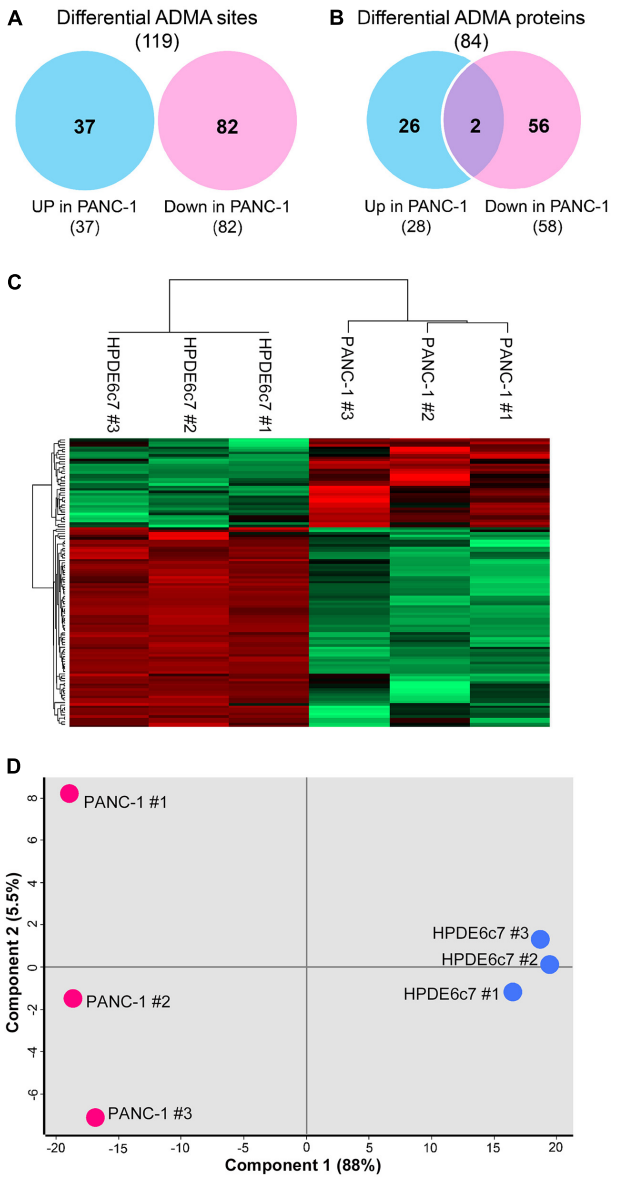


FIGURE 2 | Significantly differential ADMA profiles between HPDE6c7 and PDAC cells. **(A,B)** Numbers of ADMA sites and ADMA-containing proteins with significantly differential dimethylation levels between HPDE6c7 and PANC-1 cells. Differentially dimethylated arginine sites were defined by $P < 0.05$ using the Perseus software and shown in Venn diagrams. Totally 119 ADMA sites showing differential dimethylation between HPDE6c7 and PANC-1 cells originate from these 193 overlapping sites shown in **Figure 1D**. **(C)** Hierarchical clustering of ADMA sites showing differential dimethylation between HPDE6c7 and PANC-1 cells. The higher and lower dimethylation levels are shown by red and green colors, respectively. **(D)** The PCA evaluation of differential asymmetric arginine dimethylation between HPDE6c7 and PANC-1 cells. Biological replicates of HPDE6c7 and PANC-1 cells are shown in blue and red circles, respectively. The proportions of the variance of component 1 and 2 are shown in the brackets.

and gene transcription, especially mRNA processing and stability, RNA splicing, RNA localization and transport, chromatin organization, nucleosome assembly, and transcription initiation (**Figures 4A,B**). The separate categorization of ADMA-containing proteins in HPDE6c7 and PANC-1 cells disclosed enrichments in biological processes similar to the total ADMA-containing proteins (data not shown). In terms of molecular functions, these ADMA-containing proteins possessed pleiotropic activities as transcription factors, transcription activators or repressors, RNA polymerase II transcription factors, and DNA or mRNA binding molecules (**Figures 4A,C**). These ADMA-containing proteins were also significantly enriched in nuclear components such as nuclear lumen, nucleoplasm, nucleolus, chromosome, nuclear body, and spliceosome (**Figures 4A,D**). Furthermore, many ADMA-containing proteins were associated with signaling pathways mediated by intracellular receptors, steroid hormone receptors, and androgen receptors (**Figure 4A**), which was mainly observed in proteins with lower ADMA levels in PANC-1 cells compared with HPDE6c7 cells, but not in those showing higher ADMA levels in PANC-1 cells (**Figure 4B**).

ADMA-Containing Proteins Enriched in Spliceosome and Cancer-Related Pathways

For more insights into the possible roles of asymmetric arginine dimethylation in PDAC, we further analyzed the enrichment of ADMA-containing proteins in KEGG signaling pathways. These ADMA-harboring proteins in HPDE6c7 and PANC-1 cells were significantly associated with spliceosome, basal transcription, systemic lupus erythematosus, Notch signaling, and multiple cancer pathways including acute myeloid leukemia, chronic myeloid leukemia, and thyroid cancer (**Figure 5A** and **Supplementary Figures 1–7**). Specifically, several spliceosome components were asymmetrically dimethylated including Sm, SF3b, SR140, ACINUS, hnRNPs, and SR (**Figure 5B**). Moreover, two ADMA-containing proteins showing higher dimethylation in PANC-1 cells were related to the spliceosome pathway, suggesting the involvement of RNA splicing regulation by asymmetric arginine dimethylation during PDAC development (**Figure 5C**).

However, two other ADMA-containing proteins with lower dimethylation in PANC-1 cells were associated with the acute myeloid leukemia pathway (**Figure 5C**). Consistently, proteins dimethylated only in HPDE6c7 were enriched in cancer-related pathways such as endometrial cancer, colorectal cancer, and acute myeloid leukemia, while proteins dimethylated only in PANC-1 cells were enriched in the systemic lupus erythematosus pathway which is not directly associated with cancer pathogenesis (**Figure 5D**). Moreover, these ADMA-containing proteins were also enriched in the Wnt/β-catenin, Hedgehog, tumor growth factor beta (TGF-β), mitogen-activated protein kinase (MAPK), and other cancer-related signaling pathways (**Figure 6A**). Among them, the dimethylation of TCF, Sos, and AML in PANC-1 cells was significantly lower than the HPDE6c7 cells, while Dvl and PML proteins showed higher dimethylation in PANC-1 cells compared with HPDE6c7 cells

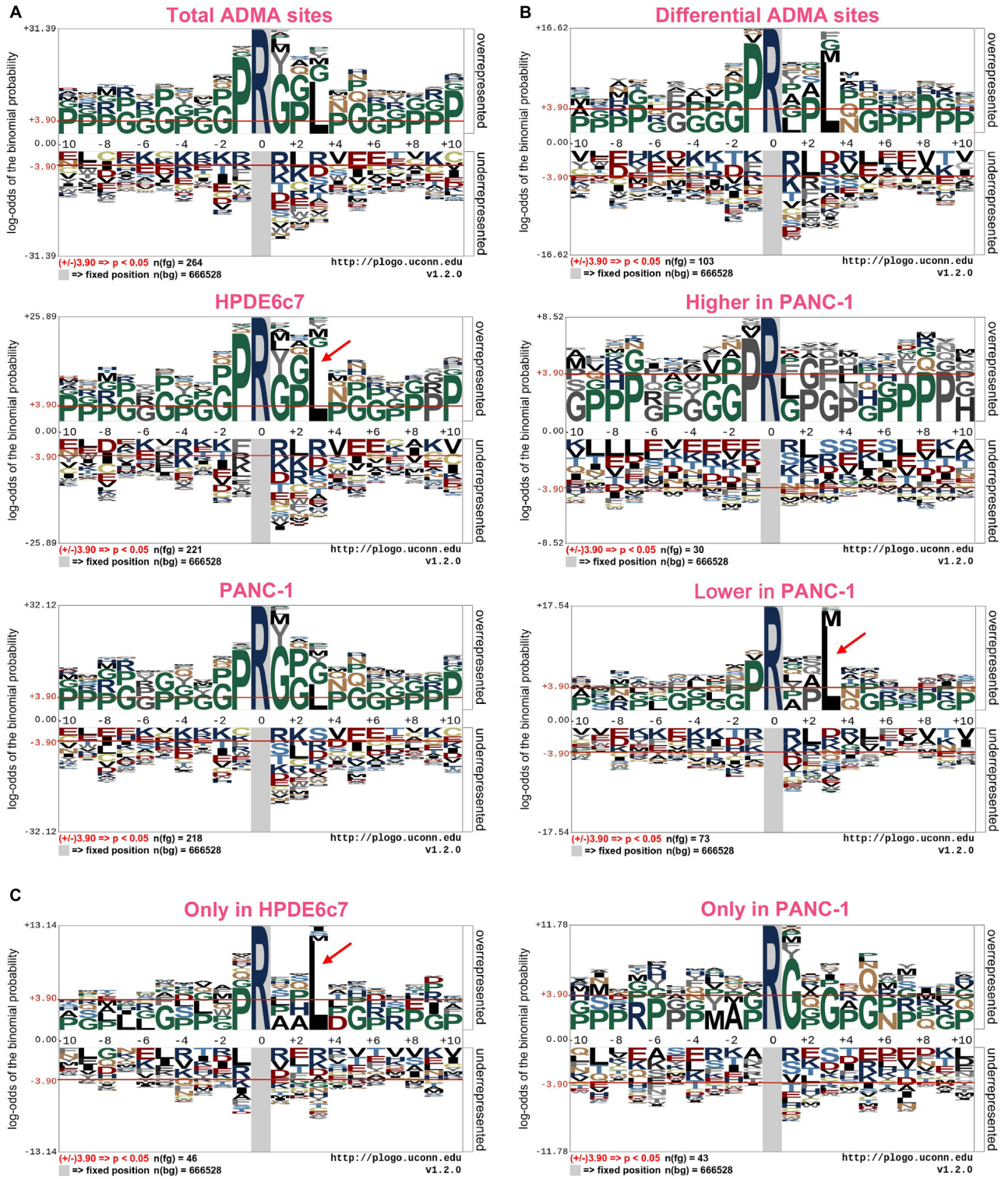


FIGURE 3 | Consensus sequences of ADMA-containing peptides in PDAC cells. **(A)** Consensus sequences of a total of 351 ADMA-containing peptides identified in HPDE6c7 and PANC-1 cells. **(B)** Consensus sequences of ADMA-containing peptides differentially dimethylated between HPDE6c7 and PANC-1 cells. The enrichment was based on 119 ADMA-containing peptides with significantly differential dimethylation between HPDE6c7 and PANC-1 cells. **(C)** Consensus sequences of ADMA-containing peptides, detected only in HPDE6c7 or PANC-1 cells. The dimethylated arginine residues are indicated in dark blue. Frequencies of amino acid residues flanking ADMA sites were analyzed using the pLogo software, and significant enrichment was defined by $P < 0.05$. The significant overrepresentation of leucine is highlighted by red arrows.

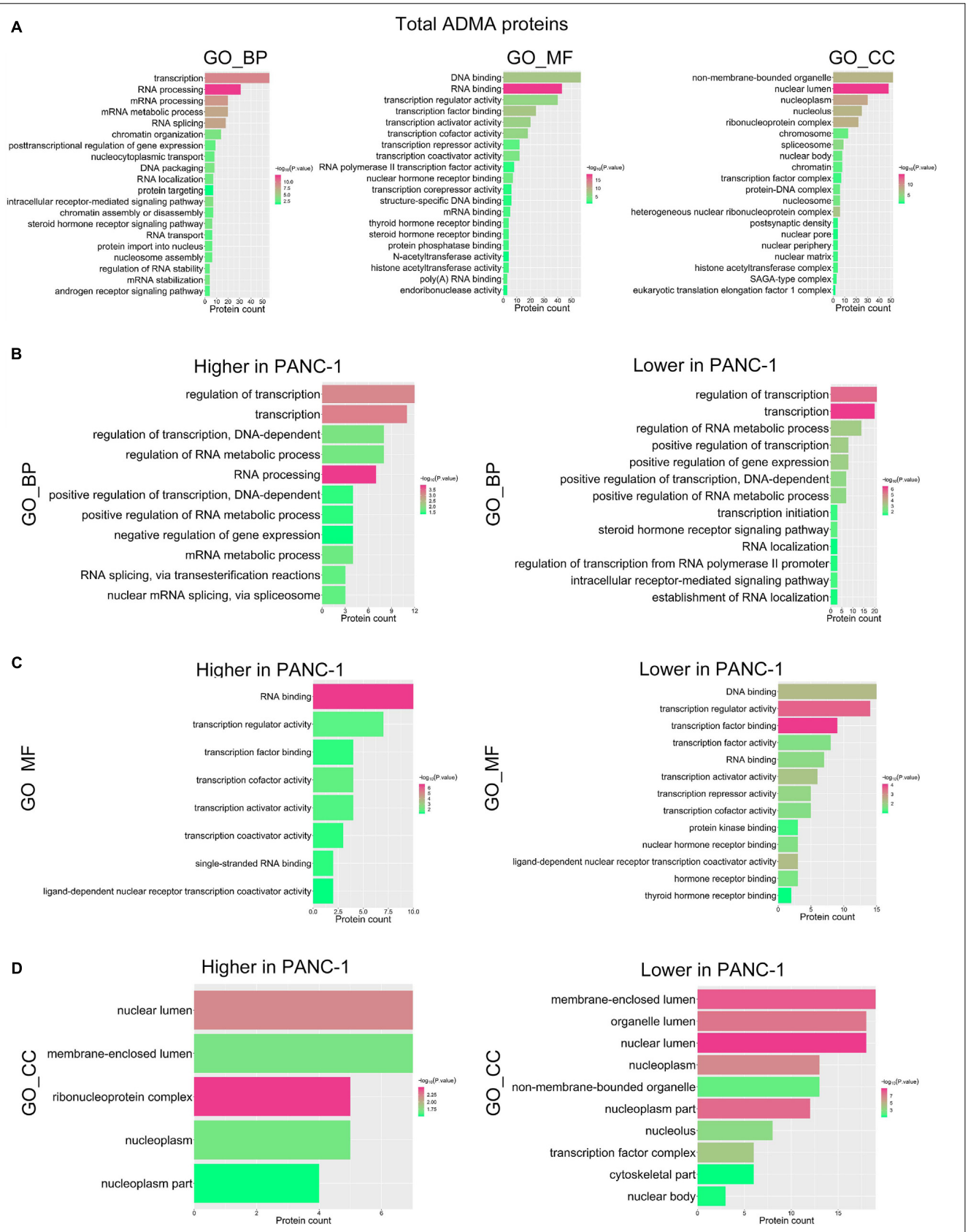


FIGURE 4 | Continued

FIGURE 4 | Functional categorization of ADMA-containing proteins in PDAC cells. **(A)** Functional annotation of total ADMA-containing proteins in HPDE6c7 and PANC-1 cells. The GO biological processes (BP), molecular functions (MF), and cellular components (CC) with significant enrichment of ADMA-containing proteins are presented as red and green bars ($P < 0.05$). **(B–D)** Functional categorization of ADMA-containing proteins showing significantly differential dimethylation between HPDE6c7 and PANC-1 cells. The classification of ADMA-containing proteins were based on GO biological processes **(B)**, molecular functions **(C)**, and cellular components **(D)**, respectively.

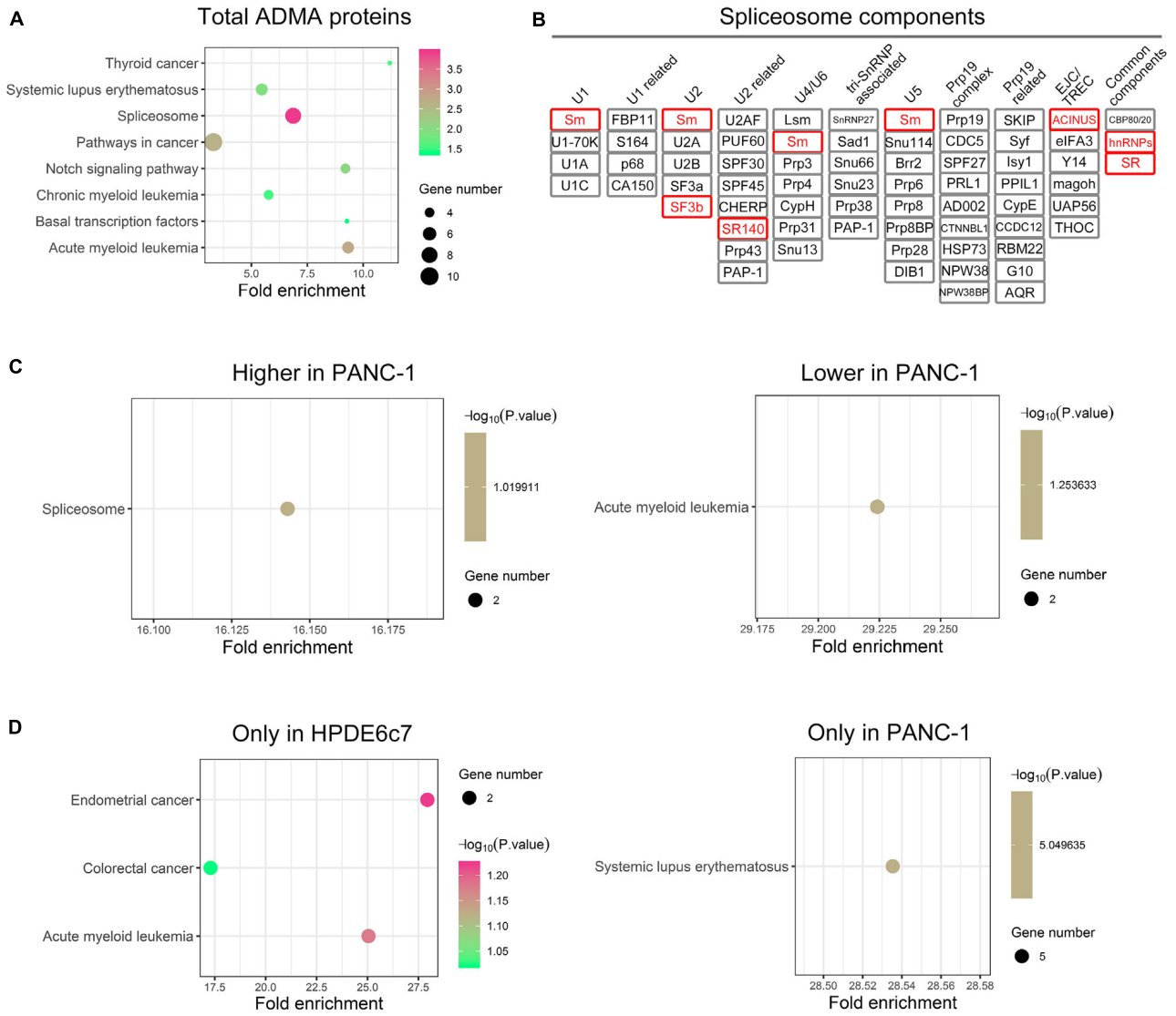


FIGURE 5 | The enrichments of ADMA-containing proteins in KEGG pathways. **(A)** Signaling pathways with significant enrichment of total ADMA-containing proteins detected in HPDE6c7 and PANC-1 cells. The enrichments in KEGG signaling pathways were analyzed by searching against the Database for Annotation, Visualization, and Integrated Discovery (DAVID). **(B)** ADMA-containing spliceosome component proteins identified in HPDE6c7 and PANC-1 cells. ADMA-containing proteins identified in this study are shown in red rectangular. **(C)** The enrichment of ADMA-containing proteins with differential dimethylation between HPDE6c7 and PANC-1 cells in KEGG signaling pathways. **(D)** KEGG signaling pathways with enrichments of ADMA-containing proteins detected exclusively in HPDE6c7 or PANC-1 cells.

(**Figure 6A**). Four other cancer-related proteins STAT5, TRK, GLI, and CtBP showed no significant dimethylation alterations between HPDE6c7 and PANC-1 cells (**Figure 6A**). The extensive dimethylation of cancer-related pathways suggested the possible roles of asymmetric arginine dimethylation in PDAC pathogenesis.

PRMT4 Expression and ADMA-Containing Protein Levels Associated With PDAC Cell Viability

PRMT4 performs essential roles in regulating pancreatic cancer cell functions by catalyzing ADMA formation (Wang Y. P. et al., 2016). In total, 103 ADMA-containing peptides which

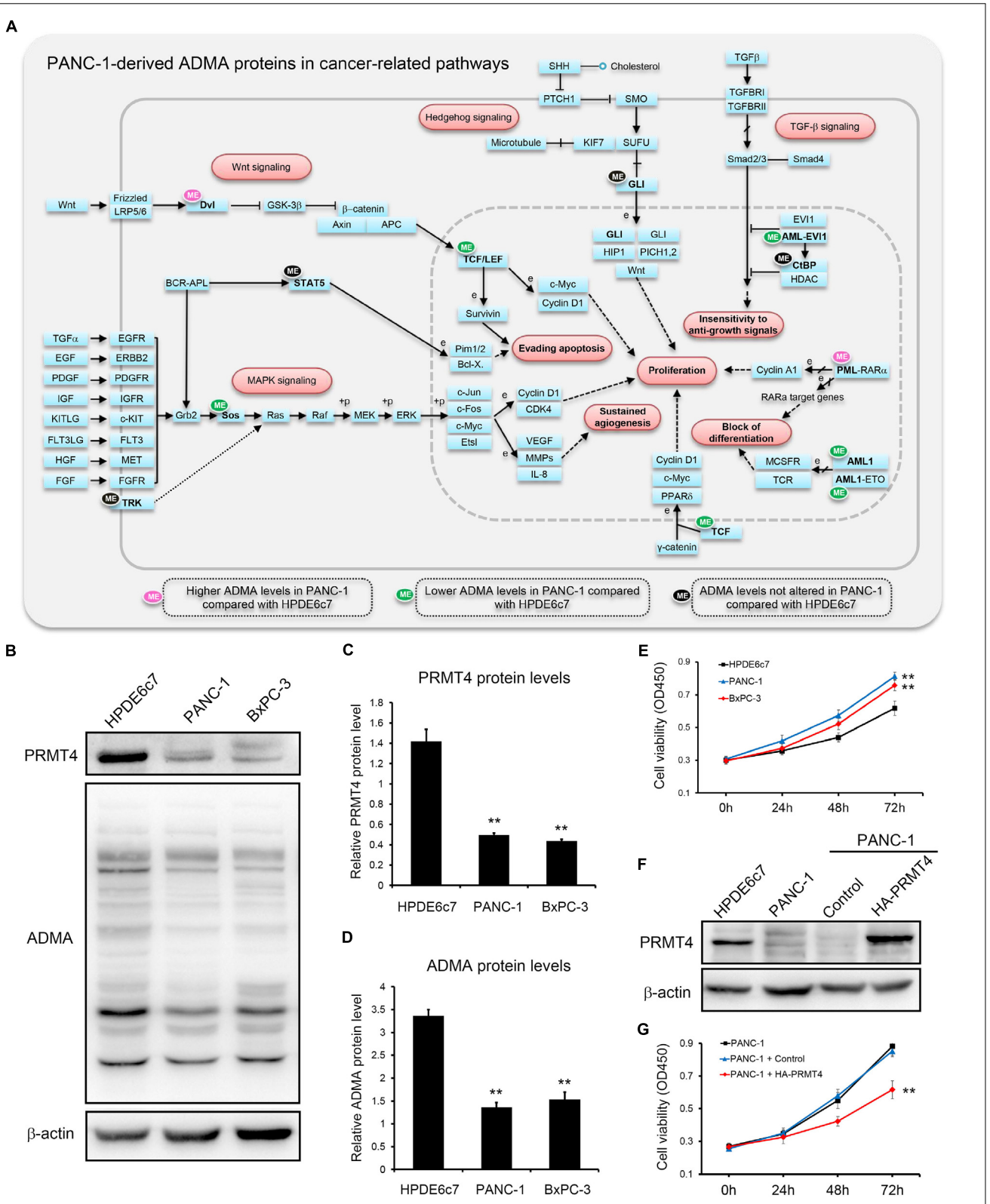


FIGURE 6 | Extensive asymmetric arginine dimethylation associated with PDAC cell viability. **(A)** ADMA-containing proteins in major cancer-related signaling pathways. Total ADMA-containing proteins were searched against the KEGG pathway database (www.kegg.jp). ADMA-containing proteins showing higher, lower, or similar asymmetric arginine dimethylation levels in PANC-1 cells by proteomics are indicated by pink, green, and black ellipses, respectively, compared with

(Continued)

FIGURE 6 | Continued

HPDE6c7 cells as the control. **(B)** Alterations of PRMT4 protein abundances and total ADMA-containing protein levels in HPDE6c7, PANC-1, and BxPC-3 cells. Western blotting was performed using β-actin as the internal standard. **(C,D)** Quantitative analysis of PRMT4 abundances and total ADMA-containing proteins levels among HPDE6c7, PANC-1, and BxPC-3 cells detected in **(B)**. Relative protein levels calibrated by β-actin from three biological replicates were presented as mean ± standard deviation and analyzed by the ANOVA method using SPSS 18.0 software. **(E)** The higher viabilities of PANC-1 and BxPC-3 cells compared with HPDE6c7 cells. Cell viabilities were determined by the CCK-8 method. The differences in cell viabilities in three biological replicates were analyzed by the ANOVA method. **(F)** Overexpression of HA-PRMT4 proteins in PANC-1 cells. The stable overexpression of recombinant HA-PRMT4 proteins in PANC-1 cells were validated by western blotting using β-actin as the internal standard. **(G)** The decrease of PANC-1 cell viability induced by PRMT4 overexpression. CCK-8 assay was carried out to measure cell viability of PRMT4-overexpressing PANC-1 cells. PANC-1 cells transfected with packaged lentivirus vectors without ligation of PANC-1 coding sequences were used as the control group. Cell viability differences based on three biological replicates were analyzed by the ANOVA method, $^{**}P < 0.01$.

were identified in our proteomic assay, including 63 peptides differentially methylated in PANC-1 cells, were previously characterized as PRMT4 substrates (Shishkova et al., 2017; **Supplementary Tables 1**, **2**). We also showed that PRMT4 protein abundances in two PDAC cell lines PANC-1 and BxPC-3 cells were significantly lower than the HPDE6c7 cells (**Figures 6B,C**). Consistently, total ADMA-containing protein levels in PANC-1 and BxPC-3 cells were substantially lower than in the HPDE6c7 cells (**Figures 6B,D**). Moreover, we found that PANC-1 and BxPC-3 cells showed significantly enhanced cell viabilities compared with HPDE6c7 cells (**Figure 6E**). To preliminarily analyze the roles of ADMA in PDAC, we overexpressed the PRMT4 gene in PANC-1 cells, which moderately increased asymmetric arginine dimethylation in only two protein bands, possibly due to limited resolution of western blotting (**Figure 6F** and **Supplementary Figure S8**). PRMT4 overexpression significantly repressed the viability of PANC-1 cells (**Figure 6G**). These results provided a basis for elucidating the roles of asymmetric arginine dimethylation in regulating PANC-1 cell functions and PDAC pathogenesis.

DISCUSSION

Asymmetric arginine dimethylation catalyzed by type I PRMTs like PRMT4 critically regulates cancer development (Blanc and Richard, 2017; Murn and Shi, 2017; Guccione and Richard, 2019; Jarrold and Davies, 2019). As well as the epigenetic regulation mediated by histone methylation, the formation of ADMA in non-histone proteins also emerged as an essential regulator of cancer pathogenesis (Biggar and Shawn, 2015; Guccione and Richard, 2019; Jarrold and Davies, 2019). However, the roles of asymmetric arginine dimethylation in PDAC remains poorly understood, mainly due to the limited ADMA-containing proteins identified in PDAC cells. In the present study, we performed proteomic profiling of ADMA-containing peptides in HPDE6c7 and PANC-1 cells. Most ADMA-containing peptides in PANC-1 cells showed significantly lower dimethylation compared with HPDE6c7 cells. Importantly, for the first time, we showed the significant differences in consensus sequences flanking the ADMA sites between HPDE6c7 and PANC-1 cells. Moreover, ADMA-harboring proteins were associated with RNA splicing and multiple cancer signaling pathways. Finally, we preliminarily validated the roles of asymmetric arginine dimethylation in regulating PANC-1 cell viability by overexpressing PRMT4. These results suggest the essential roles of asymmetric arginine dimethylation in PDAC cells, which provide a basis for the functional elucidation of ADMA-containing proteins in PDAC pathogenesis.

In recent years, immunoaffinity purification using antibodies targeting methylated arginine residues has been successfully applied in the proteomic characterization of protein methylation (Fisk et al., 2013; Guo et al., 2014; Sylvestersen et al., 2014; Geoghegan et al., 2015; Shishkova et al., 2017; Yakubu et al., 2017; Zeeshan et al., 2017), but not in pancreatic cancer. Here, we reported the first global characterization of ADMA-containing proteins in pancreatic cancer, which identified 289 ADMA sites among over 200 proteins by label-free quantitative proteomics combined with affinity purification. Among them, the asymmetric arginine dimethylation of 102 proteins were previously reported, indicative of the reliability of the proteomic dataset. For instance, TAF15, a nuclear RNA-binding protein involved in gene transcription and formation of fusion oncoprotein, was asymmetrically dimethylated at Arg206 to enhance target gene expression (Jobert et al., 2009). Its asymmetric dimethylation at Arg206 was also identified in our proteomic analysis. The methylation of MED12, which sensitizes breast cancer cells to chemotherapy (Wang et al., 2015), was also detected in pancreatic cancer cells. Meanwhile, 99 ADMA sites were newly identified in this study and further investigations might provide new insights into arginine methylation-mediated PDAC development. Our quantitative analysis also showed differential dimethylation of 119 ADMA sites in PANC-1 cells, which needs further experimental validation via calibration by their total protein levels.

As indicated in previous reports, proline and glycine residues were significantly enriched in the flanking sequences of ADMA sites, which were termed as the proline-, glycine-, and methionine-rich (PGM) motif or the glycine and arginine-rich (GAR) motif (Uhlmann et al., 2012; Fisk et al., 2013; Guo et al., 2014; Shishkova et al., 2017). Specifically, glycine residues near ADMA sites were predicted to facilitate the access of arginine residue to PRMT active sites by enhancing conformational flexibility (Blanc and Richard, 2017). Meanwhile, proline residues in the vicinity of ADMA sites might promote ADMA formation by enhancing substrate recognition and increasing hydrophobicity (Shishkova et al., 2017). The significant enrichment of the proline and glycine residues flanking ADMA sites was also observed here in HPDE6c7 and PANC-1 cells, which indicated the unbiased ADMA-containing peptide characterization in this proteomic study. Importantly, we also demonstrated significant enrichment of the leucine

residues flanking ADMA sites in pancreatic cells, suggesting the special biochemical mechanism of ADMA formation in pancreatic biology. Considering the hydrophobic property of leucine residues, this observation indicates that higher hydrophobicity might be required for asymmetric arginine dimethylation in pancreatic cells. Furthermore, significant leucine enrichment was observed in ADMA-containing peptides with higher dimethylation or only detected in HPDE6c7 cells, other than those with higher dimethylation or only detected in PANC-1 cells. The biochemical mechanisms and the functional significance of reduced leucine enrichment in pancreatic cancer cells deserves further investigation.

Aberrant RNA splicing contributes to cancer development by inducing transcriptome changes and signaling alterations (Dong and Chen, 2019). Protein arginine methylation also performs key roles in regulating RNA splicing events (Deng et al., 2010; Sanchez et al., 2010; Hu et al., 2017; Jarrold and Davies, 2019). In this methylome study, six spliceosome components were identified in pancreatic cancer cells, further supporting the roles of arginine methylation in regulating RNA splicing during cancer pathogenesis. Moreover, multiple components of cancer signaling pathways were identified in our proteomic assay, such as the Wnt/β-catenin, Hedgehog, TGF-β and MAPK pathways. For instance, transcription factor 7-like 2 (TCF7L2) is one transcriptional partner of the Wnt/β-catenin pathway regulating aerobic glycolysis in pancreatic cancer (Xiang et al., 2018), which showed significantly lower dimethylation in PANC-1 cells in our quantitative proteomics. Moreover, STAT5 protein inactivation mediated Cucurbitacin B-induced G2-M-phase arrest and apoptosis in pancreatic cancer cells (Thoennissen et al., 2009). This study showed that STAT5 protein was also asymmetrically dimethylated in PDAC cells, and its potential roles in pancreatic cancer cell responses to chemotherapies are still worth exploring. In addition, the dimethylated proteins detected only in HPDE6c7 cells were significantly enriched in cancer-related pathways (**Figure 5D**), which suggested the suppression of their dimethylation in PANC-1 cells. Further investigation is required to broaden knowledge of protein methylation-mediated cancer signaling during PDAC pathogenesis.

PRMT4 performs critical roles in cancer biology by catalyzing asymmetric arginine dimethylation and its expression is repressed in pancreatic cancer cells (Wang Y. P. et al., 2016). The decrease of PRMT4 expression and lower ADMA-containing protein levels were also detected in our study, in two PDAC cell lines compared with the normal pancreatic epithelial cells. Importantly, PRMT4 overexpression caused partial recovery of asymmetric arginine dimethylation and significant suppression of PANC-1 cell viability, which was consistent with a previous report showing the involvement of PRMT4 catalytic activity in regulating PANC-1 cell growth (Wang Y. P. et al., 2016). Moreover, nearly half of the dimethylated proteins identified in HPDE6c7 and PANC-1 cells were previously characterized as PRMT4 substrates by a comprehensive PRMT4 methylome study in breast cancer cells (Shishkova et al., 2017). The ADMA sites identified in this study could serve as clues for further elucidation of the PDAC pathogenic mechanisms mediated by asymmetric arginine dimethylation. In contrast to its high expression in breast cancer (Wang et al., 2014), the greatly repressed expression of PRMT4 in PDAC cells also suggested that the pathogenic roles of PRMT4 in pancreatic cancer might be mediated by tissue-specific signaling mechanisms. In addition, other type I PRMTs such as PRMT1 also critically regulate pancreas development and PDAC pathogenesis (Wang Y. et al., 2016; Lin et al., 2018; Lee et al., 2019; Song et al., 2020), and their pathogenic roles in PDAC cells deserve further investigations.

In summary, we reported the proteomic characterization of ADMA-containing protein profiles between HPDE6c7 and PANC-1 cells in this study. The majority of ADMA-containing proteins showed significantly lowered dimethylation in PANC-1 cells, with a reduced leucine residue enrichment flanking ADMA sites. These ADMA-containing proteins were associated with spliceosome machinery and multiple cancer signaling pathways. Moreover, PRMT4 overexpression partially recovered asymmetric arginine dimethylation and repressed viability in PANC-1 cells. These results provided a comprehensive view of asymmetric arginine dimethylation profiles in PDAC cells, which would facilitate elucidating the pathogenic roles of protein arginine methylation in pancreatic cancer.

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

JH and TC conceived and designed the study. MW performed the main cellular and proteomic assays, assisted by ZT, YL, YC, and CC. CT, YH, TC, and JH undertook the data analysis and bioinformatics. JH, TC, CT, and WZ wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (grant number 81600502), the Natural Science Foundation of Hunan Province (grant number 2019JJ40164), the Guangzhou University of Chinese Medicine (grant number A1-AFD018181Z3943), the Hunan Provincial Education Department (grant number 19B352), and the Young Doctor Training Program of Hunan Provincial People's Hospital (grant number BSJJ202004).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.545934/full#supplementary-material

Supplementary Figure 1 | ADMA-containing proteins in the spliceosome machinery. A schematic illustration of ADMA-containing proteins significantly enriched in the spliceosome machinery. The diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 2 | ADMA-containing proteins in the basal transcription processes. A schematic illustration of ADMA-containing proteins significantly enriched in the basal transcription processes. The diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 3 | ADMA-containing proteins in the acute myeloid leukemia pathway. A schematic illustration of ADMA-containing proteins significantly enriched in the acute myeloid leukemia pathway. The signaling pathway diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 4 | ADMA-containing proteins in the chronic myeloid leukemia pathway. A schematic illustration of ADMA-containing proteins significantly enriched in the chronic myeloid leukemia pathway. The signaling pathway diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 5 | ADMA-containing proteins in the thyroid cancer pathway. A schematic illustration of ADMA-containing proteins significantly enriched in the thyroid cancer pathway. The signaling pathway diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 6 | ADMA-containing proteins in the Notch signaling pathway. A schematic illustration of ADMA-containing proteins significantly enriched in the Notch signaling pathway. The signaling pathway diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 7 | ADMA-containing proteins in the systemic lupus erythematosus pathway. A schematic illustration of ADMA-containing proteins significantly enriched in the systemic lupus erythematosus pathway. The signaling pathway diagram was modified from KEGG pathway database (www.kegg.jp).

Supplementary Figure 8 | Total ADMA-containing protein levels in PRMT4-overexpressing PANC-1 cells. Total ADMA-containing protein levels were detected by western blotting and protein bands showing the increases of ADMA-containing proteins induced by PRMT4 overexpression were indicated by red arrows.

Supplementary Table 1 | A list of total ADMA-containing peptides identified in HPDE6c7 and PANC-1 cells. The protein accessions, names, methylation sites, Andromeda scores, peptide intensities and previous reports of ADMA-containing peptides identified in HPDE6c7 and PANC-1 cells were listed in this table.

Supplementary Table 2 | A list of ADMA-containing peptides showing differential dimethylation between HPDE6c7 and PANC-1 cells. Information of ADMA-containing peptides with significantly differential dimethylation between HPDE6c7 and PANC-1 cells were analyzed by the Perseus software and listed in this table, including protein accessions, names, sequence windows, logarithm-transformed intensities and *P*-values from the Student *T*-tests.

REFERENCES

Bedford, M. T., and Clarke, S. G. (2009). Protein arginine methylation in mammals. Who, What, and Why. *Mol. Cell* 33, 1–13. doi: 10.1016/j.molcel.2008.12.013

Biggar, K. K., and Shawn, S.-C. L. (2015). Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* 16:5. doi: 10.1038/nrm3915

Blanc, R. S., and Richard, S. (2017). Arginine methylation. The coming of age. *Mol. Cell* 65:8. doi: 10.1016/j.molcel.2016.11.003

Buscail, L., Bournet, B., and Cordelier, P. (2020). Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. *Nat. Rev. Gastroenterol. Hepatol.* 17, 153–168. doi: 10.1038/s41575-019-0245-4

Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer statistics in China, 2015. *CA Cancer J. Clin.* 66, 115–132. doi: 10.3322/caac.21338

Cheung, N., Fung, T. K., Zeisig, B. B., Holmes, K., Rane, J. K., Mowen, K. A., et al. (2016). Targeting aberrant epigenetic networks mediated by PRMT1 and KDM4C in acute myeloid leukemia. *Cancer Cell* 29, 32–48. doi: 10.1016/j.ccell.2015.12.007

Deng, X., Gu, L., Liu, C., Lu, T., Lu, F., Lu, Z., et al. (2010). Arginine methylation mediated by the *Arabidopsis* homolog of PRMT5 is essential for proper pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19114–19119. doi: 10.1073/pnas.1009669107

Dong, X., and Chen, R. (2019). Understanding aberrant RNA splicing to facilitate cancer diagnosis and therapy. *Oncogene* 39, 2231–2242. doi: 10.1038/s41388-019-1138-2

Fisk, J. C., Li, J., Wang, H., Aletta, J. M., Qu, J., and Read, L. K. (2013). Proteomic analysis reveals diverse classes of arginine methylproteins in mitochondria of trypanosomes. *Mol. Cell. Proteom. MCP* 12:302. doi: 10.1074/mcp.m112.022533

Geoghegan, V., Guo, A., Trudgian, D., Thomas, B., and Acuto, O. (2015). Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. *Nat. Commun.* 6:6758.

Guccione, E., and Richard, S. (2019). The regulation, functions and clinical relevance of arginine methylation. *Nat. Rev. Mol. Cell Biol.* 20, 642–657. doi: 10.1038/s41580-019-0155-x

Guo, A., Gu, H., Zhou, J., Mulhern, D., Wang, Y., Lee, K. A., et al. (2014). Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell. Proteom.* 13, 372–387. doi: 10.1074/mcp.o113.027870

Hsu, M. C., Pan, M. R., Chu, P. Y., Tsai, Y. L., Tsai, C. H., Shan, Y. S., et al. (2018). Protein arginine methyltransferase 3 enhances chemoresistance in pancreatic cancer by methylating hnRNPA1 to increase ABCG2 expression. *Cancers* 11:8. doi: 10.3390/cancers11010008

Hu, J., Huang, X., Chen, L., Sun, X., Lu, C., Zhang, L., et al. (2015). Site-specific nitrosoproteomic identification of endogenously S-nitrosylated proteins in *Arabidopsis*. *Plant Physiol.* 167, 1731–1746. doi: 10.1104/pp.15.00026

Hu, J., Yang, H., Mu, J., Lu, T., Peng, J., Deng, X., et al. (2017). Nitric oxide regulates protein methylation during stress responses in plants. *Mol. Cell* 67, 702–710. doi: 10.1016/j.molcel.2017.06.031

Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Iovanna, J., and Dusetti, N. (2017). Speeding towards individualized treatment for pancreatic cancer by taking an alternative road. *Cancer Lett.* 410, 63–67. doi: 10.1016/j.canlet.2017.09.016

Jarrold, J., and Davies, C. C. (2019). PRMTs and arginine methylation. Cancer's best-kept secret? *Trends Mol. Med.* 25, 993–1009. doi: 10.1016/j.molmed.2019.05.007

Jobert, L., Argentini, M., and Tora, L. (2009). PRMT1 mediated methylation of TAF15 is required for its positive gene regulatory function. *Exp. Cell Res.* 315, 1273–1286. doi: 10.1016/j.yexcr.2008.12.008

Lee, K., Kim, H., Lee, J., Oh, C. M., Song, H., Kim, H., et al. (2019). Essential role of protein arginine methyltransferase 1 in pancreas development by regulating protein stability of neurogenin 3. *Diabetes Metab. J.* 43, 649–658. doi: 10.4093/dmj.2018.0232

Lin, Z., Chen, Y., Lin, Z., Chen, C., and Dong, Y. (2018). Overexpressing PRMT1 inhibits proliferation and invasion in pancreatic cancer by inverse correlation of ZEB1. *IUBMB Life* 70, 1032–1039. doi: 10.1002/iub.1917

Liu, F., Ma, F., Wang, Y., Hao, L., Zeng, H., Jia, C., et al. (2017). PKM2 methylation by CARM1 activates aerobic glycolysis to promote tumorigenesis. *Nat. Cell Biol.* 19:1358. doi: 10.1038/ncb3630

Ma, J., Chen, T., Wu, S., Yang, C., Bai, M., Shu, K., et al. (2019). iProX. an integrated proteome resource. *Nucleic Acids Res.* 47, D1211–D1217.

Murn, J., and Shi, Y. (2017). The winding path of protein methylation research. milestones and new frontiers. *Nat. Rev. Mol. Cell Biol.* 18:517. doi: 10.1038/nrm.2017.35

O'Shea, J. P., Chou, M. F., Quader, S. A., Ryan, J. K., Church, G. M., and Schwartz, D. (2013). pLogo. a probabilistic approach to visualizing sequence motifs. *Nat. Methods* 10, 1211–1212. doi: 10.1038/nmeth.2646

Pan, S., Brentnall, T. A., and Chen, R. (2020). Proteome alterations in pancreatic ductal adenocarcinoma. *Cancer Lett.* 469, 429–436. doi: 10.1016/j.canlet.2019.11.020

Roth, G. S., Casanova, A. G., Lemonnier, N., and Reynoird, N. (2017). Lysine methylation signaling in pancreatic cancer. *Curr. Opin. Oncol.* 30, 30–37. doi: 10.1097/cco.0000000000000421

Ryan, D. P., Hong, T. S., and Bardeesy, N. (2014). Pancreatic adenocarcinoma. *N. Engl. J. Med.* 371, 1039–1049.

Sanchez, S. E., Petrillo, E., Beckwith, E. J., Zhang, X., Rugnone, M. L., Hernando, C. E., et al. (2010). A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature* 468, 112–116. doi: 10.1038/nature09470

Shishkova, E., Zeng, H., Liu, F., Kwiecien, N. W., Hebert, A. S., Coon, J. J., et al. (2017). Global mapping of CARM1 substrates defines enzyme specificity and substrate recognition. *Nat. Commun.* 8:15571.

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.* 68, 7–30. doi: 10.3322/caac.21442

Song, C., Chen, T. W., He, L., Ma, N., Li, J. A., Rong, Y. F., et al. (2020). PRMT1 promotes pancreatic cancer growth and predicts poor prognosis. *Cell. Oncol.* 43, 51–62. doi: 10.1007/s13402-019-00435-1

Sylvestersen, K. B., Horn, H., Jungmichel, S., Jensen, L. J., and Nielsen, M. L. (2014). Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest. *Mol. Cell. Proteom.* 13, 2072–2088. doi: 10.1074/mcp.o113.032748

Tan, C., Li, Y., Huang, X., Wei, M., Huang, Y., Tang, Z., et al. (2019). Extensive protein S-nitrosylation associated with human pancreatic ductal adenocarcinoma pathogenesis. *Cell Death Dis.* 10:914.

Thoennissen, N. H., Iwanski, G. B., Doan, N. B., Okamoto, R., Lin, P., Abbassi, S., et al. (2009). Cucurbitacin B induces apoptosis by inhibition of the JAK/STAT pathway and potentiates antiproliferative effects of gemcitabine on pancreatic cancer cells. *Cancer Res.* 69, 5876–5884. doi: 10.1158/0008-5472.can-09-0536

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., et al. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740. doi: 10.1038/nmeth.3901

Uhlmann, T., Geoghegan, V. L., Thomas, B., Ridlova, G., Trudgian, D. C., and Acuto, O. (2012). A method for large-scale identification of protein arginine methylation. *Mol. Cell. Proteom.* 11, 1489–1499. doi: 10.1074/mcp.m112.020743

Waldmann, T., and Schneider, R. (2013). Targeting histone modifications–epigenetics in cancer. *Curr. Opin. Cell Biol.* 25, 184–189. doi: 10.1016/j.ceb.2013.01.001

Wang, L., Zeng, H., Wang, Q., Zhao, Z., Boyer, T. G., Bian, X., et al. (2015). MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs. *Sci. Adv.* 1:e1500463. doi: 10.1126/sciadv.1500463

Wang, L., Zhao, Z., Meyer, M. B., Saha, S., Yu, M., Guo, A., et al. (2014). CARM1 methylates chromatin remodeling factor BAF155 to enhance tumor progression and metastasis. *Cancer Cell* 25, 21–36. doi: 10.1016/j.ccr.2013.12.007

Wang, Y., Hsu, J. M., Kang, Y., Wei, Y., Lee, P. C., Chang, S. J., et al. (2016). Oncogenic functions of Gli1 in pancreatic adenocarcinoma are supported by its PRMT1-mediated Methylation. *Cancer Res.* 76:7049. doi: 10.1158/0008-5472.can-16-0715

Wang, Y. P., Zhou, W., Wang, J., Huang, X., Zuo, Y., Wang, T. S., et al. (2016). Arginine methylation of MDH1 by CARM1 inhibits glutamine metabolism and suppresses pancreatic cancer. *Mol. Cell.* 64:673. doi: 10.1016/j.molcel.2016.09.028

Wu, J., Cui, N., Wang, R., Li, J., and Wong, J. (2012). A role for CARM1-mediated histone H3 arginine methylation in protecting histone acetylation by releasing corepressors from chromatin. *PLoS One* 7:e34692. doi: 10.1371/journal.pone.0034692

Xiang, J., Hu, Q., Qin, Y., Ji, S., Xu, W., Liu, W., et al. (2018). TCF7L2 positively regulates aerobic glycolysis via the EGLN2/HIF-1alpha axis and indicates prognosis in pancreatic cancer. *Cell Death Dis.* 9:321.

Yakubu, R. R., Silmon de Monerri, N. C., Nieves, E., Kim, K., and Weiss, L. M. (2017). Comparative monomethylarginine proteomics suggests that protein arginine methyltransferase 1 (PRMT1) is a significant contributor to arginine monomethylation in *Toxoplasma gondii*. *Mol. Cell. Proteom.* 16, 567–580. doi: 10.1074/mcp.m117.066951

Yang, Y., and Bedford, M. T. (2013). Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* 13, 37–50. doi: 10.1038/nrc3409

Zeeshan, M., Kaur, I., Joy, J., Saini, E., Paul, G., Kaushik, A., et al. (2017). Proteomic identification and analysis of arginine-methylated proteins of *Plasmodium falciparum* at asexual blood stages. *J. Proteome Res.* 16, 368–383. doi: 10.1021/acs.jproteome.5b01052

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 © 2020 Wei, Tan, Tang, Lian, Huang, Chen, Chen, Zhou, Cai and Hu. 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 Oncology

ORIGINAL RESEARCH
published: 14 December 2020
doi: 10.3389/fonc.2020.550213

Comparative Analysis of Systemic and Tumor Microenvironment Proteomes From Children With B-Cell Acute Lymphocytic Leukemia at Diagnosis and After Induction Treatment

Geise Ellen Broto[1,2†], Stephany Corrêa[3†], Fausto Celso Trigo[4], Everton Cruz dos Santos[3], Fernanda Tomiotto-Pelissier[5], Wander Rogério Pavanelli[5], Guilherme Ferreira Silveira[6], Eliana Abdelhay[3†] and Carolina Panis[1,2,5,7†]*

[1] Programa de Pós-graduação em Patologia Clínica e Laboratorial, Universidade Estadual de Londrina, Londrina, Brazil, [2] Laboratório de Biologia de Tumores, Universidade Estadual do Oeste do Paraná, UNIOESTE, Francisco Beltrão, Brazil, [3] Laboratório de Células-Tronco, Centro de Transplante de Medula Óssea (CEMO), Instituto Nacional de Câncer, Rio de Janeiro, Brazil, [4] Pediatrics Department, Hospital do Câncer de Londrina, Londrina, Brazil, [5] Programa de Pós-graduação em Patologia Experimental Universidade Estadual de Londrina, Londrina, Brazil, [6] Fundação Osvaldo Cruz, Instituto Carlos Chagas, Curitiba, Brazil, [7] Programa de Pós-Graduação em Ciências Aplicadas à Saúde, Universidade Estadual do Oeste do Paraná, UNIOESTE, Francisco Beltrão, Brazil

Among the childhood diseases, B-cell acute lymphocytic leukemia (B-ALL) is the most frequent type of cancer. Despite recent advances concerning disease treatment, cytotoxic chemotherapy remains the first line of treatment in several countries, and the modifications induced by such drugs in the organism are still poorly understood. In this context, the present study provided a comparative high-throughput proteomic analysis of the cumulative changes induced by chemotherapeutic drugs used in the induction phase of B-ALL treatment in both peripheral blood (PB) and bone marrow compartment (BM) samples. To reach this goal, PB and BM plasma samples were comparatively analyzed by using label-free proteomics at two endpoints: at diagnosis (D0) and the end of the cumulative induction phase treatment (D28). Proteomic data was available *via* ProteomeXchange with identifier PXD021584. The resulting differentially expressed proteins were explored by bioinformatics approaches aiming to identify the main gene ontology processes, pathways, and transcription factors altered by chemotherapy, as well as to understand B-ALL biology in each compartment at D0. At D0, PB was characterized as a pro-inflammatory environment, with the involvement of several downregulated coagulation proteins as KNG, plasmin, and plasminogen. D28 was characterized predominantly by immune response-related processes and the super expression of the transcription factor IRF3 and transthyretin. RUNX1 was pointed out as a common transcription factor found in both D0 and D28. We chose to validate the proteins transthyretin and interferon-gamma (IFN-γ) by commercial kits and expressed the

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Marilia Seelaender, University of São Paulo, Brazil
Alex Kentsis, Cornell University, United States

***Correspondence:**
Carolina Panis
carolpanis@hotmail.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *08 April 2020*
Accepted: *06 November 2020*
Published: *14 December 2020*

Citation:
Broto GE, Corrêa S, Trigo FC, Santos ECd, Tomiotto-Pelissier F, Pavanelli WR, Silveira GF, Abdelhay E and Panis C (2020) Comparative Analysis of Systemic and Tumor Microenvironment Proteomes From Children With B-Cell Acute Lymphocytic Leukemia at Diagnosis and After Induction Treatment. Front. Oncol. 10:550213. doi: 10.3389/fonc.2020.550213

results as PB/BM ratios. Transthyretin ratio was augmented after induction chemotherapy, while IFN-γ was reduced at the end of the treatment. Considering that most of these proteins were not yet described in B-ALL literature, these findings added to understanding disease biology at diagnosis and highlighted a possible role for transthyretin and IFN-γ as mechanisms related to disease resolution.

Keywords: interferon-gamma, interferon regulatory factor 3, transthyretin, tumor biology, proteomics, chemotherapy, B-acute lymphocytic leukemia

INTRODUCTION

B-cell acute lymphocytic leukemia (B-ALL) is the most frequent neoplasia in childhood worldwide (1). Regardless of the great success in B-ALL treatment, this disease consists of an important cause of death in the child population. Considerable advances have been reached in understanding the biology of leukemia, and most are due to the scientific community efforts to bring new information based on high-throughput studies based on molecular approaches, that have driven recently B-ALL treatment to powerful targeted treatments (2).

The clonal rise of leukemia cells is mainly associated with cumulative mutations that directly affect the panel of proteins that are secreted in the tumor microenvironment (3). Some of these proteins are responsible for cellular events in the leukemic niche that control disease spread (4) and can result from systemic transcriptional changes related to disease prognosis (5). Drug therapy can also have a profound impact in B-ALL biology, further than killing its malignant clones. Changes in the local and systemic immunological profiles have been reported in B-ALL patients after completing conventional chemotherapy, and are related to their immune recovering (6). Remodeling induced by treatment in the genomic profile of blood cells from B-ALL has been reported as a determinant of prognosis and associated with disease outcomes as relapse and death (7).

Recently, proteomics-based approaches have been revealed as valuable tools to map changes in the protein profiling in both blood and bone marrow of B-ALL patients (8), and are helping to explore gaps related to disease relapse (9), chemoresistance (10), and biomarker discovery (11). Despite this, studies focusing on understanding the comparative analysis between the bone marrow tumor microenvironment and the peripheral blood are still missing. In this context, some overarching challenges as the hidden mechanisms behind immature lymphoid cells accumulation in the bone marrow, as well the mechanisms underlying the chemotherapy effects against B-ALL need to be improved, and therefore, the study of leukemia biology and the cumulative impact of the initial phase of chemotherapy become necessary.

For this purpose, this study performed the proteomic analysis of samples from children diagnosed with B-ALL by comparing bone marrow and peripheral blood profiles, before chemotherapy starting - at diagnosis, and the cumulative effects found after the induction treatment. To reach this goal, we used the nano-ultra performance liquid chromatography label-free proteomic strategy to obtain the differential proteomic profiles and further investigated the putative mechanistic evidence by using bioinformatics tools.

METHODS

Patients Selection and Sample Collection

A total of 17 children diagnosed with B-ALL attended from September 2014 to January 2016 in the Londrina Cancer Hospital, Londrina-Paraná, Brazil, were enrolled. Those responsible for the children signed their consent terms. This study was approved by the Institutional Ethics Committee (approval number CAAE 24498213.0.0000.5231) and was designed and conducted following the ethical principles for medical research involving human subjects from the Declaration of Helsinki. The Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) criteria were followed regarding patient selection, assay performance, and data analysis throughout the study. All samples were routinely screened in the laboratory to determine the type of leukemia, and only those diagnosed as B-ALL were included in the study.

All patients presented as high risk at diagnosis, and were submitted to the same drug schedule, as determined by the Brazilian Group for Treatment of Childhood Leukemia scheme - GBTLI 2009, displayed in **Figure 1**. The induction protocol was based on a combined schedule based on prednisone, L-asparaginase, vincristine, and doxorubicin. Whole peripheral blood and bone marrow aspirate samples were collected in EDTA tubes from recently diagnosed patients, before starting any treatment (D0) and at the end of the chemotherapy induction phase (D28). Samples were centrifuged at 4,000 rpm for 5 min at 4^0C to the obtention of plasma and kept frozen until analysis. Each sample was supplemented with a 1,000:1 (µl) protease inhibitor cocktail (GE Healthcare, USA).

Label-Free Protein Quantitation *via* Mass-Spectrometry (MS)

Proteomic analysis was conducted with pooled plasma samples for each B-ALL group. Protein quantification was obtained *via* the Bradford assay, and the samples were processed using Amicon columns device of 3 kDa ultra-filtration (Millipore, USA) for concentration (39x) and exchanged buffer with 50 mM NH_4HCO_3. A total of 200 µg of protein were used for subsequent treatment with Rapigest 2 µg/µl (Waters), DTT 100 mM (Sigma), IAA 100 mM (Sigma), and tryptic digestion 0.25 µg/µl (High-grade Trypsin, Promega), carried out at 37°C,

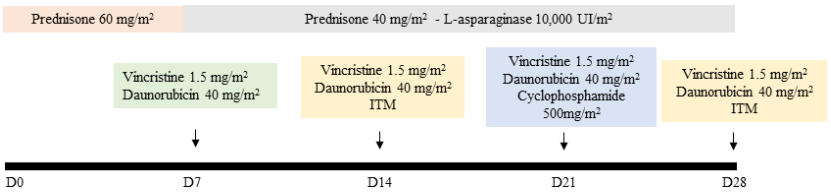


FIGURE 1 | Induction phase chemotherapy schedule - Prednisone 60 mg/m^2 orally from D1 to D7, then 40 mg/m^2/day orally, divided into 2–3 doses for 3 weeks (D8-29), suspending regressively in 3 to 4 days. If necessary, Prednisolone can be administered intravenously, divided into 3 doses. Vincristine: 1.5 mg/m^2/week intravenously maximum dose of 2 mg) administered on days 8, 15, 22, and 29. Daunorubicin: 40 mg/m^2/week intravenously, administered on days 8, 15, and 22. L-Asparaginase: 10,000 IU/m^2 intramuscular or intravenously (if thrombocytopenia <75,000/mm3) every 3 days, from day 8 of treatment, for a total of 9 doses. Cyclophosphamide: 500 mg/m^2 intravenously on days 22 and 23 of induction for patients in the slow response subgroup. Intrathecal medication (ITM): triple therapy with Methotrexate, Ara-C and Dexamethasone will be administered at age-adjusted doses, on days 15 and 29 of induction (> 1 <3 years: 10mg/m^2 and 20mg/m^2, for Methotrexate and Ara-C respectively; > 3 <9 years: 12 mg/m^2 and 24 mg/m^2, respectively; > 9 years: 15 mg/m^2 and 30 mg/m^2, respectively. The dose of Dexamethasone is uniform (2 mg/m^2. Max. 2 mg dose). Reference: Brazilian Group for Treatment of Childhood Leukemia scheme – GBTLI 2009 (12).

overnight and under light agitation (400 rpm). The proteomic approach applied in this study was the nano-Ultra Performance Liquid Chromatography (nano-UPLC) tandem nano-ESI-HDMSE method for qualitative and quantitative experiments. A nanoACQUITY UPLC system (Waters, UK) was employed, as previously reported by our group (13, 14). Briefly, a strong cation exchange column (180 µm × 23 mm, Waters, England) packed with Polysulfoethyl aspartamide (5 µm, PolyLC, USA) was used for the first dimension. Nine salt gradient fractions were used to elute the samples from the strong cation exchange column, followed by an RP gradient. After the peptides were captured, the trap column was placed online with a different RP analytical column (100 µm × 100 mm, 1.8 µm C18, nanoACQUITY UPLC HSS T3, Waters, UK), and an RP gradient of 5%–40% acetonitrile (containing 0.1% v/v formic acid) in 58 min was used as the second dimension, with a flow rate of 600 nL min−1. Analyses were performed using nano-ESI in positive ion mode [nanoESI (+)] with a NanoLockSpray ionization source (Waters, UK). Multiplexed data-independent scanning with specificity and selectivity based on nonlinear "T-wave" ion mobility (HDMSE) experiments were performed with a Synapt HDMS mass spectrometer (Waters, UK), as previously described (15). Full scan orthogonal acceleration TOF (oa-TOF) MSE data were acquired from m/z 50 to 2,000.

Database Searching, Quantification, and Statistical Analysis

Database searching and protein quantification were performed as previously reported (16–18). Using the PLGS Expression E tool algorithm, the identified proteins were organized into a statistically significant list corresponding to increased and decreased regulation ratios between the plasma from peripheral blood compared to plasma from bone marrow at D0 and the plasma from peripheral blood compared to plasma from bone marrow at D28. *In silico* analysis for biological processes, canonical pathways, network interactions, and transcription factors were performed using Metacore™ software (Clarivate Analytics, https://portal.genego.com/).

Validation Study

Transthyretin and interferon-gamma (IFN-γ) were chosen as targets for the validation step. Transthyretin (also known as pre-albumin) was measured in PB and BM plasma by a commercial kit based on immune turbidimetry assay (Aptec Diagnostics, Belgium). Plasma samples were individually diluted 1:10 in a saline buffer and an aliquot of 50 µl was added to 900 µl of reaction buffer. This mixture was read at 340 nm (OD_1). Then, 60 µl of goat anti-human pre-albumin antibody was added to this mixture and read again at 340 nm after a 5-min incubation (OD_2). The difference between both OD was obtained, and transthyretin levels were calculated against a standard calibration curve. Interferon-gamma levels were determined in samples by using the Human Th1/Th2 cytokine kit (BD Biosciences, catalog number 550749) by flow cytometry. A plasma aliquot of 50 µl from each patient was mixed to 50 µl of capture beads provided by the kit and incubated for 3 h with 50 µl of PE detection reagent. After each tube was washed and centrifuged, the bead pellet recovered, plated, and read after incubation in a flow cytometer (BD Accuri™). The same procedure was performed to the calibration standard curve. Results were expressed individually and as PB/BM ratios, and compared by t-test in the software GraphPad Prism 6.0, considering as significant a p-value < 0.05.

RESULTS

Table 1 shows the clinicopathological characteristics of patients. A total of 17 B-ALL children were enrolled in the study, the mean age at diagnosis was 6.8 years, 9 patients were male, most of them were Caucasian, and the mean body mass index was 16.8 kg/m^2. The mean leukocyte counting in the PB at D0 was 18,273 cells/mm^3,ranging from 200 to 74,300 white blood cells/mm^3, and 1,609 cells/mm^3 at D28 in PB, ranging from 200 to 9,700 white blood cells/mm^3. The mean tumor cells counting in BM aspirate at D0 was 56%, ranging from 28%–90%, and 6.5% at D28, ranging from 0%–11%.

TABLE 1 | Clinicopathological data of B-ALL patients.

	B-ALL patients
Total number of patients	n = 17
Gender	n = 8
Female	n = 9
Male	
Ethnicity	n = 15
Caucasian	n = 2
Afro-descendant	
Mean age at diagnosis, years (min-max)	6.9 (1.7–15)
Mean body mass index, kg/m^2 (min-max)	16.8 (13.13–25)
Mean leukocyte counting, peripheral blood cells/ mm^3 (min-max)	18,273 (200–74,300)
D0	1,609 (200–9700)
D28	
Mean blast counting in bone marrow aspirate, percentage (min-max)	56 (28–90)
D0	6,5 (0–41)
D28	

Figure 2 shows the results of the initial high-throughput screening of blood and bone marrow samples at D0 and D28. As shown in **Figure 2A**, 91 proteins were initially identified as differentially expressed, with 7 proteins expressed uniquely in the blood, 21 down expressed in blood, and 18 exclusively expressed in the bone marrow. At D28, 94 proteins were initially identified as differentially expressed among blood and bone marrow, being 8 overexpressed in blood, while 12 were down expressed in blood being exclusive from bone marrow. A false discovery rate (FDR) of a maximum of 4% was applied, and an FDR rate lower than 1% was detected for all analyses, at protein and peptide level, on average. Moreover, a minimum of 15 peptides, on average, was applied for protein identification.

The Venn diagram (**Figure 2B**) shows that 23 proteins were exclusively differentially expressed in the blood at D0, and 15 were exclusively differentially expressed in the blood at D28. Moreover, 5 proteins were common to both D0 and D28 (vitronectin, apolipoprotein C III, serum amyloid A2 protein, inter-alpha-trypsin inhibitor heavy chain H4 and putative protein ZNF720, **Figure 2C**). Vitronectin was upregulated in both comparisons whereas apolipoprotein C III and putative protein ZNF720 were downregulated in both comparisons. However, 2 proteins had their expression shifted among comparisons: serum amyloid A2 protein and inter-alpha-trypsin inhibitor heavy chain H4, which indicate their putative association in treatment response.

Box 1 shows the final list containing differentially identified proteins after 2-fold change cut-off and statistical significance, and their status as up or downregulated in the blood at D0 and D28. The full list of identified proteins, together with raw data and details regarding methods were deposited in ProteomeXChange repository (Project accession: PXD021584, Project DOI: 10.6019/PXD021584).

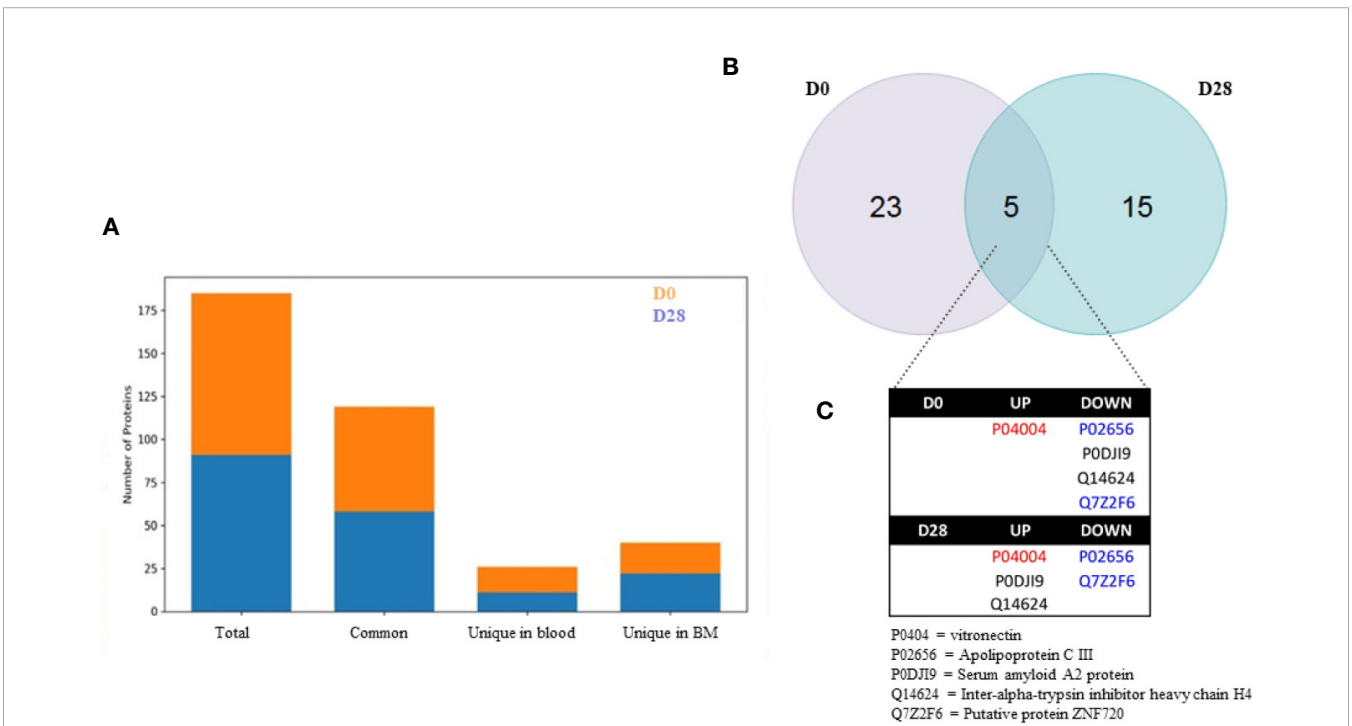


FIGURE 2 | Proteomic design of the study **(A)** and the number of common and unique proteins identified in the high-throughput proteomic screening. In **(B)**, Venn's diagram from the differentially expressed proteins in our study. Up and downregulated proteins differentially expressed in peripheral blood plasma samples from ALL-B patients at D0 – at diagnosis, before treatment start, and D28 – at the end of the treatment, comparing the blood versus the bone marrow. **(C)** Highlight from the 5 protein differentially expressed among D0 and D28. In red: upregulated proteins in both comparisons; In blue: downregulated proteins in both comparisons; In black: proteins in which quantification shifted according to comparison to bone marrow. * Chemotherapic protocol according to the Brazilian Group for Treatment of Childhood Leukemia scheme (2009).

BOX 1 | Up and downregulated proteins differentially expressed in peripheral blood plasma samples from ALL-B patients at D0 – at diagnosis, before treatment start, and D28 – at the end of the treatment compared to bone marrow.

UPREGULATED IN BLOOD – D0

Uncharacterized protein C16orf90
Apolipoprotein C-I
Vitronectin
Inter-alpha-trypsin inhibitor heavy chain
Forkhead box protein E3
Dehydrogenase/reductase SDR family member 11
Protein RUFY3

DOWNREGULATED IN BLOOD – D0

Apolipoprotein A-II
Immunoglobulin kappa variable 3D-15
Immunoglobulin kappa variable 3-15
Potassium channel subfamily K member 1
Putative PIN1-like protein
P antigen family member 1
Plasminogen
Kininogen-1
Hemoglobin subunit epsilon
Apolipoprotein C-III
Leucine-rich alpha-2-glycoprotein
Serum amyloid A-2 protein
Notch homolog 2 N-terminal-like protein C
Hemoglobin subunit gamma-1
Hemoglobin subunit gamma-2
Paired box protein Pax-2
Complement factor H-related protein 3
Inter-alpha-trypsin inhibitor heavy chain H4
Putative protein ZNF720
Notch homolog 2 N-terminal-like protein A
Protein eva-1 homolog A

UPREGULATED IN BLOOD – D28

Transthyretin
Vitronectin
Serum amyloid A-2 protein
Inter-alpha-trypsin inhibitor heavy chain H4
Interferon regulatory factor 3
Reticulocalbin-3
Regulator of microtubule dynamics protein 1
Protein C10

DOWNREGULATED IN BLOOD – D28

Homeobox protein VENTX
Prothrombin
Apolipoprotein C-III
Homeobox protein Hox-B1
Inorganic pyrophosphatase
ATPase family AAA domain-containing protein 3B
Docking protein 3
Putative protein ZNF720
Ribonuclease 8
Zinc finger and SCAN domain-containing protein 10
WW domain-binding protein 1-like
Hsp70-binding protein 1

Figure 3 shows the significant gene ontology cellular processes and networks regarding the differentially expressed proteins in the blood at D0 and D28, as identified by using the Metacore software. At D0, the main processes identified were related to the cellular redox status and detoxification, and the networks were related to inflammation and blood coagulation signaling. At D28, the processes were related mainly to inflammation and immunity, and the networks represented mostly extracellular remodeling and innate immune response-related pathways.

The *in silico* analysis of the differentially expressed proteins at D0 in the blood of patients (**Figure 4**) revealed that some downregulated proteins, as KNG, plasmin, bradykinin, and kallidin, were enrolled in the protein folding and maturation process (**Figure 4A**), which affects inflammation. In **Figure 4B**, downregulated proteins from the coagulation system, plasminogen, KNG, and plasmin are shown. Plasminogen and plasmin are also highlighted in the map shown in **Figure 4C**, as participants of the wound-healing, proliferation, and migration processes. **Figure 4D** shows the involvement of the upregulated vitronectin in the extracellular matrix remodeling and actin cytoskeleton reorganization processes, altogether with the downregulated plasminogen and plasmin.

Figure 5 shows the main biological networks for the differentially expressed proteins found in the blood at D28. **Figure 5A** shows the enrollment of the upregulated transcription factor named interferon regulatory factor 3 (IRF3) in triggering the TLR-inflammatory cascade. In 5B, thrombin downregulation and its putative role in cell migration and tumor progression pathways. **Figure 5C** demonstrates the upregulated protein transthyretin and its participation in the stem cell differentiation process.

Metacore analysis also provided the potential transcription factors (TFs) identified as upstream regulators in ALL-B identified proteins. As shown in **Box 2**, it was observed differences between the top 10 upstream regulators pointed for D0 and D28 in the blood. Some TFs presented a difference in the score (e.g., FOXP3) and some TFs were only associated with D0 data (e.g., TAL1) or D28 data (e.g., GCR). We selected TAL1 from D0 analysis and AML1(RUNX1) from D28 analysis to visualize their regulation in the attempt to better understand the potential regulation of these TFs (**Figures 6A, B**). We observed a crosslink with other TFs, which is different among analysis (D0 x D28), together with an activation/inhibition regulation.

For validation, we chose the proteins transthyretin and interferon-gamma (a protein originated from IRF3 activation), since both represent new information in B-ALL biology. Our data showed that the PB/BM ratio for IFN-γ was reduced at D28 (1.2 ± 0.1 at D0 and 0.65 ± 0.05 at D28, **Figure 7A**, $p < 0.05$), while transthyretin was augmented at D28 (0.85 ± 0.05 at D0 and 1.25 ± 0.02 at D28, **Figure 7B**, $p < 0.05$). Individual values for both transthyretin and IFN-γ measurements in PB and BM, at D0 and D28, are displayed in 7C (means ± standard errors of the means and min-max values). Finally, the individual comparison of data shown a significant augment of IFN-γ in BM at D28 ($p < 0.05$).

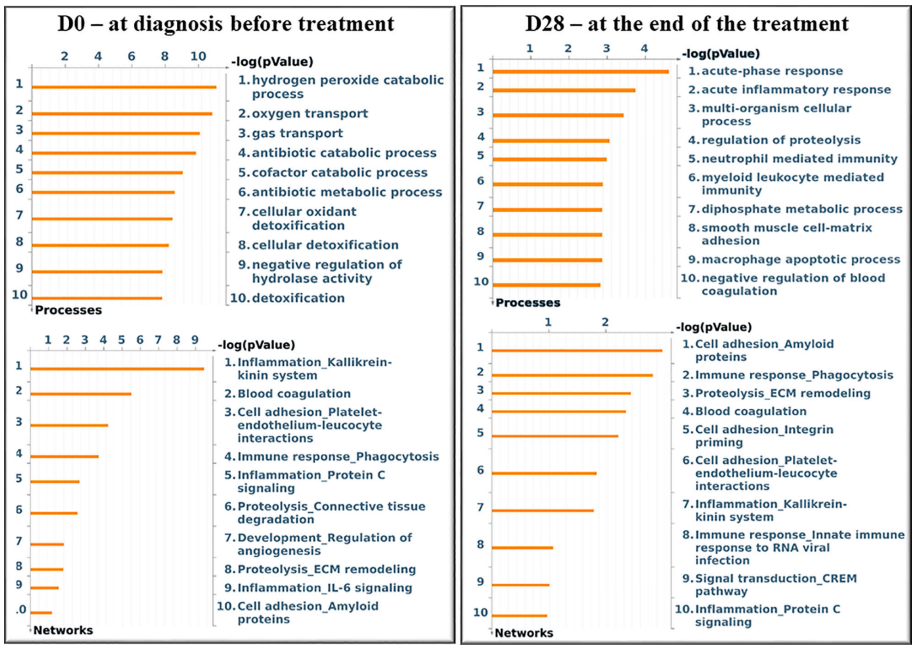


FIGURE 3 | Significant Gene Ontology cellular processes and networks identified by Metacore analysis of the differentially identified proteins from each group. *In silico* analysis was performed using GeneGO Metacore™ software (GeneGO Inc., USA). Sorting is done for the 'Statistically significant Processes.

To reinforce the relationship between the proteins discovered, we performed a word cloud study (**Figure 8**). The group of proteins present in each time point and at different levels of expression returned frequencies of different terms in a search in PubMed. Four searches were carried out with the term leukemia and proteins in each analyzed condition (A - D0 down, B - D28 down, C - D0 up, and D - D28 up), returning the 10 most recent works. The abstracts of the works were concatenated and organized as a term corpus (analyzed by the Natural Language Toolkit package - NLTK 3.4 and Word Cloud 1.6.0 for Python). The size of the word represents the frequency in the corpus and highlighted the interferon axis at the D28.

DISCUSSION

Proteomic-based strategies are powerful tools to identify new information in tumor biology studies, by using designs that are of relevance to clinical practice (19, 20). In the present study, the comparative analysis of the proteomic profile between the systemic (blood) and tumor (bone marrow) microenvironments provided a picture regarding the main proteins and processes that are present at diagnosis(D0) and triggered by the induction chemotherapy (D28) in the blood from ALL-B patients. The comparison performed between D0 and D28 allowed to describe the cumulative effect of cytotoxic treatment, understand the main processes present in ALL-B at diagnosis, and to know the main proteins differentially expressed that are relevant in each compartment before and after chemotherapy.

At diagnosis, before any treatment starting (D0), the blood was characterized as a pro-inflammatory environment, where some redox processes might be occurring. Further, it was identified that proteins that are classically viewed as coagulation players can also participate in different signaling pathways and processes in B-ALL. The involvement of coagulation proteins as KNG, plasmin, and plasminogen, that were found in our study as downregulated in the blood at diagnosis (D0), was suggested by *in silico* analysis as far beyond than clotting. The participation of such proteins in pivotal processes as wound healing and extracellular matrix remodeling has not yet been discussed in B-ALL biology.

The four main processes highlighted by the *in silico* analysis at D0 are cross-linked and display the systemic movement for tissue remodeling headed by coagulation proteins in the context of B-ALL. Protein folding processes refer to the tissue remodeling-driven events triggered by the exposure of tissue factors, in which coagulation cascade plays a crucial role. The extracellular matrix (ECM) is a broadly dynamic tissue structure that changes to allow processes as cell proliferation, migration, and differentiation (21). Therefore, at B-ALL diagnosis, it seems to exist a battle between pro and anti-ECM remodeling, headed by different proteins.

The downregulation of the coagulation proteins identified here could represent a tentative protective mechanism since the inhibition of wound healing processes could result in reduced cell proliferation and impair migration, both necessary in leukemia biology. It is known that dysregulation of ECM composition can critically support cancer progression and the

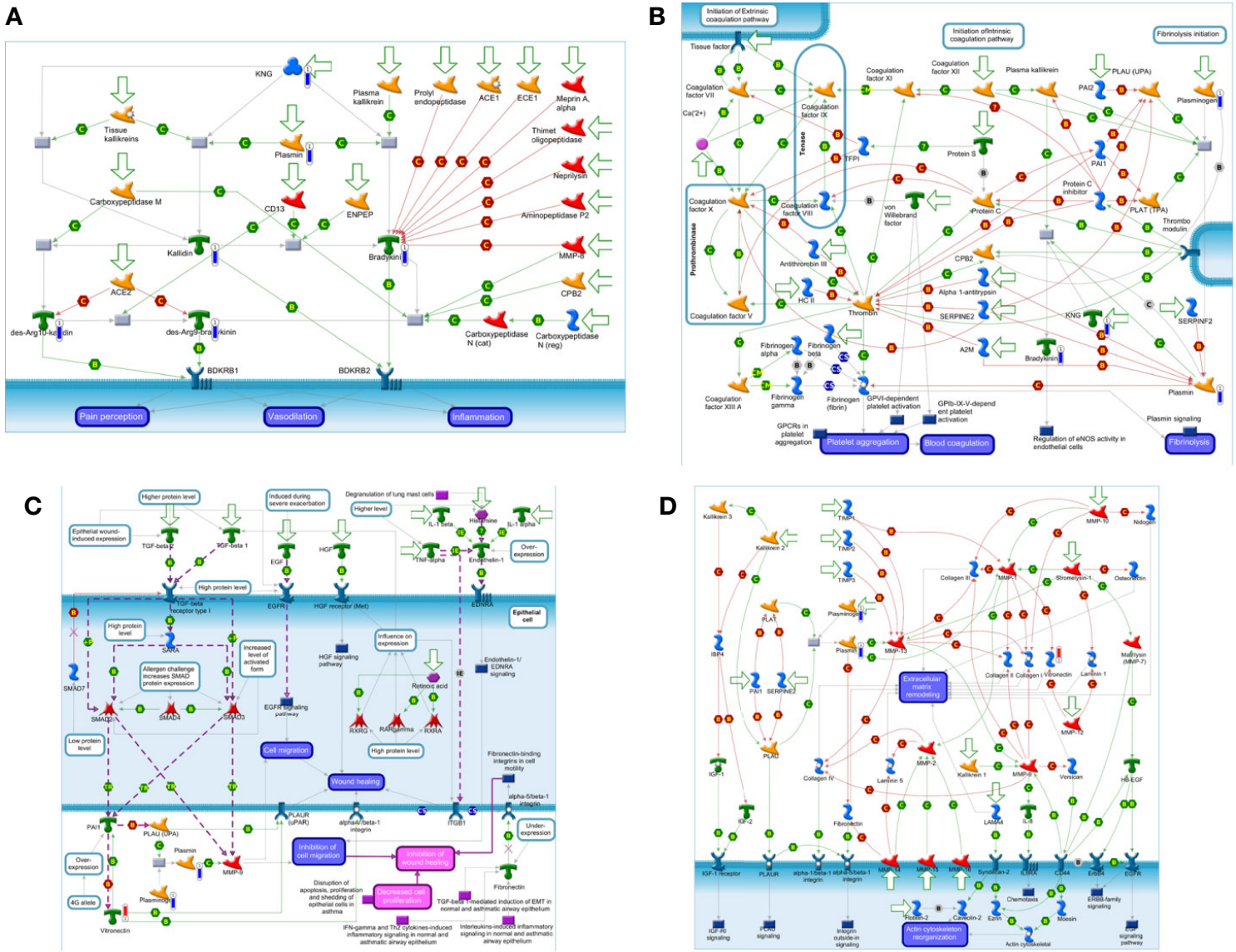


FIGURE 4 | *In silico* analysis reveals the main biological networks differentially expressed in the blood of ALL-B compared to the bone marrow at diagnosis, before treatment start (D0). **(A)** Protein folding and maturation, **(B)** Blood coagulation, **(C)** Wound healing, and **(D)** extracellular matrix remodeling. Experimental data from all files is linked to and visualized on the maps as thermometer like figures. Upward thermometers have red color and indicate up-regulated signals and downward (blue) ones indicate downregulated expression levels of the genes.

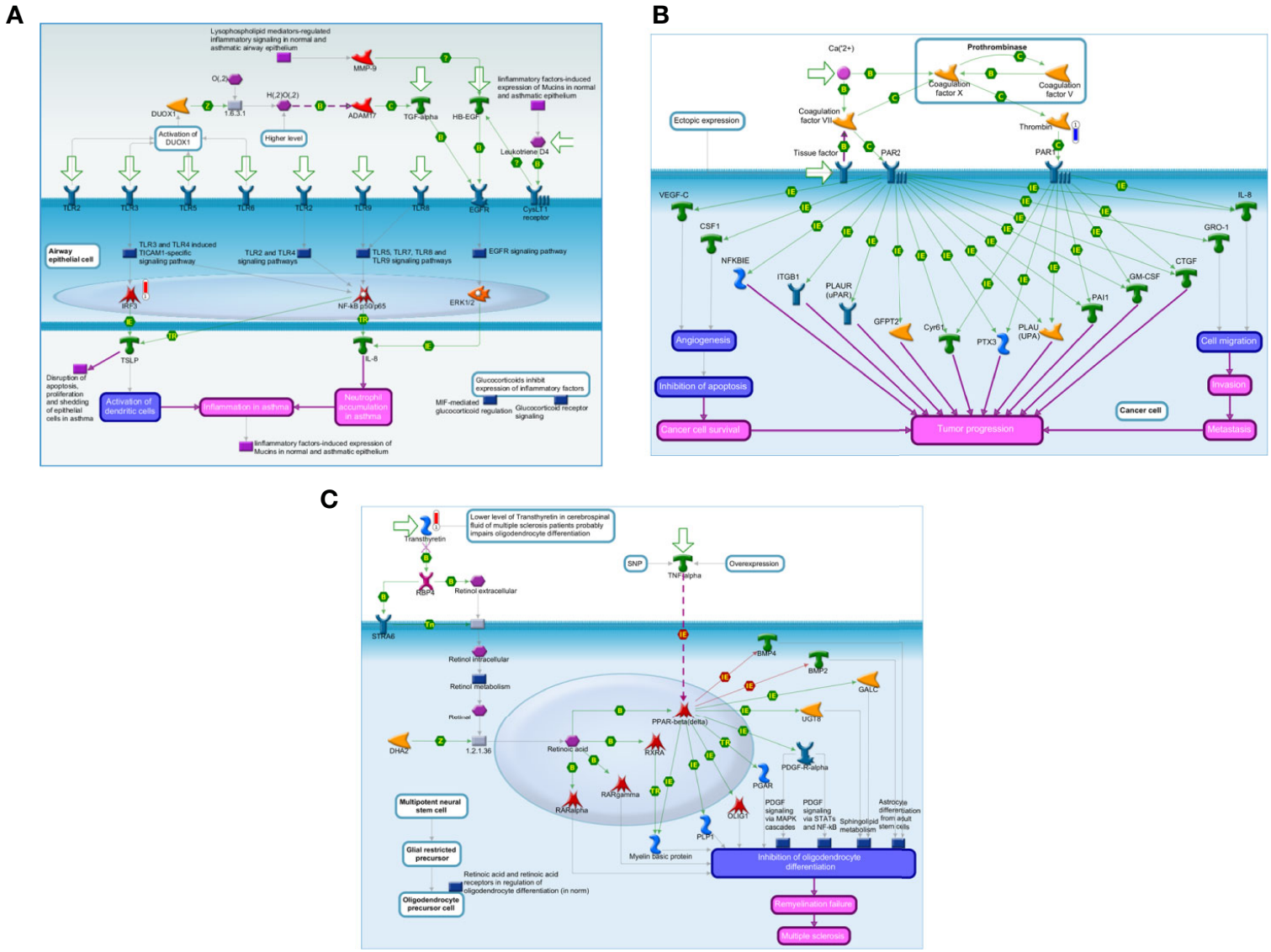


FIGURE 5 | *In silico* analysis reveals the main biological networks differentially expressed in the blood of ALL-B compared to bone marrow in D28, at the end of treatment start. **(A)** TLR and EGFR-induced inflammatory signaling, **(B)** Expression targets of Tissue factor signaling in cancer, and **(C)** 7-Retinoic acid regulation of cell differentiation. Experimental data from all files are linked to and visualized on the maps as thermometer like figures. Upward thermometers have red color and indicate upregulated signals and downward (blue) ones indicate downregulated expression levels of the genes.

BOX 2 | Top 10 upstream regulators from D0 and D28, after comparison between peripheral blood and bone marrow, identified by Metacore analysis.

DO	D28
GATA1	GATA1
TAL1	AML1 (RUNX1)
SOX17	SOX17
HNF4-alpha	HNF4-alpha
FOXP3	ETS1
AML1 (RUNX1)	GCR
SOX2	GABP-alpha
GATA2	SOX2
c-MYC	FOXP3
GLIS3	Oct-3/4

GATA1, GATA Binding Protein 1; TAL1, T-cell acute lymphocytic leukemia protein 1; SOX17, SRY-related HMG-box 17 protein; HNF4-alpha, Hepatocyte nuclear factor 4 alpha; FOXP3, forkhead box P3 or scurfin; AML1 (RUNX1), Runt-related transcription factor 1 (RUNX1) also known as acute myeloid leukemia 1 protein (AML1) or core-binding factor subunit alpha-2; SOX2, SRY-Box Transcription Factor 2; GATA2, GATA-binding factor 2; c-MYC, c-MYC proto-oncogene; Basic Helix-Loop-Helix transcription factor (bHLH); GLIS3, GLIS Family Zinc Finger 3; ETS1, v-ets erythroblastosis virus E26 oncogene homolog 1; GCR, glucocorticoid receptor; GABP-alpha, GA-binding protein alpha chain; Oct-3/4, Octamer binding transcription factor 3/4.

malignant behavior of the cells by changing the interplay between the tumor microenvironment and the cells who reside there (22). Bradykinin, for example, can have its availability affected negatively by metalloproteinases (23), while the inhibition of plasminogen activation attenuates the metastatic behavior of breast cancer cells (24). Therefore, the downregulation of these proteins could be positive for tumor establishment.

On the other hand, vitronectin seems to act here as a pro-ECM protein, since its up-regulation could affect positively the reorganization of the cytoskeleton by matrix metalloproteinases (25), favoring cell organization to spread. On cancer stem cells, it has been shown that vitronectin is the component present in human serum that drives stem cell differentiation through an integrin-dependent mechanism, responsible for tumor formation (26).

At the end of the induction phase (D28), which represents the cumulative effect of all drugs administered in patients for B-ALL shutdown, we found that the main processes and networks were related to the immune-mediated inflammatory response. We have previously reported that chemotherapy can regulate negatively the systemic acute-phase proteins in breast cancer patients, as coagulation proteins, by using the same analytical strategy (27). The downregulation of thrombin found in our

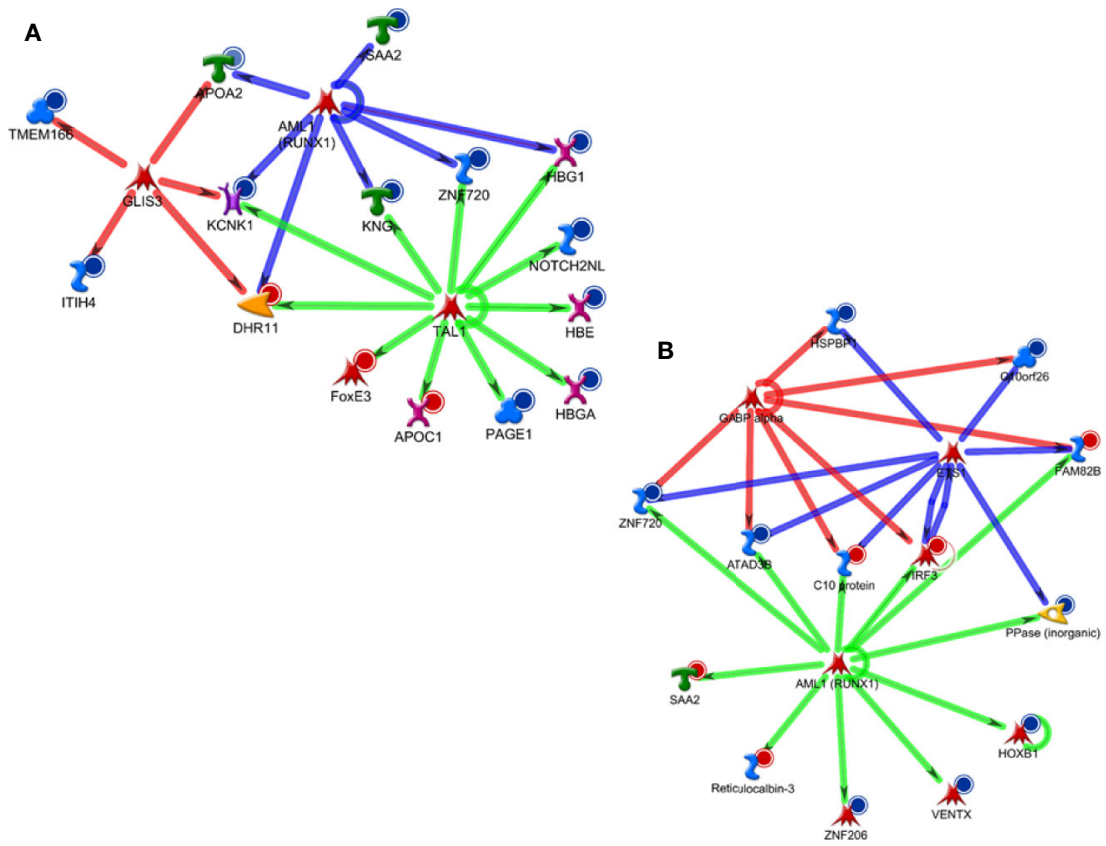
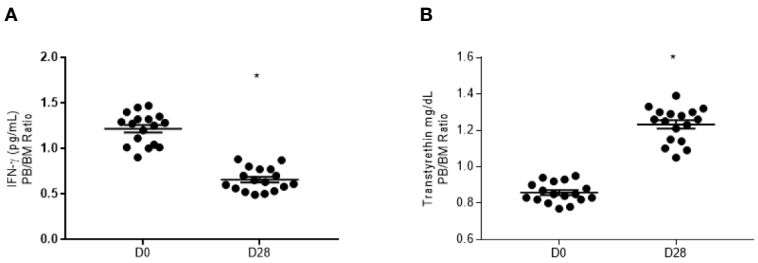


FIGURE 6 | Major potential upstream regulators based on differentially expressed proteins at the blood in D0 **(A)** and D28 **(B)**.



C

	Transthyretin mg/dL	IFN-γ pg/mL
PB – D0	12.7 ± 0.9 (11.1-13.5)	13.1 ± 3.6 (10-16.1)
BM – D0	14.45 ± 2.4 (15.9-12.0)	11.0 ± 2.1 (8.8-14.1)
PB – D28	13.9 ± 0.6 (12.9-14.6)	11.7 ± 3.6 (8.6-15.5)
BM – D28	11.1 ± 1.2 (10.3-12.7)	20.1 ± 5.5 (26-14.1)

$p < 0.05$

FIGURE 7 | Validation of transthyretin and interferon-gamma (IFN-γ) levels in plasma samples from B-ALL patients. Peripheral blood (PB)/bone marrow (BM) ratio for **(A)** interferon-gamma and **(B)** transthyretin levels. In **(C)**, specific levels for each protein at diagnosis (D0) and the end of the treatment (D28), represented as mean ± standard error of the mean (min-max). * indicates statistical significance, $p < 0.05$.

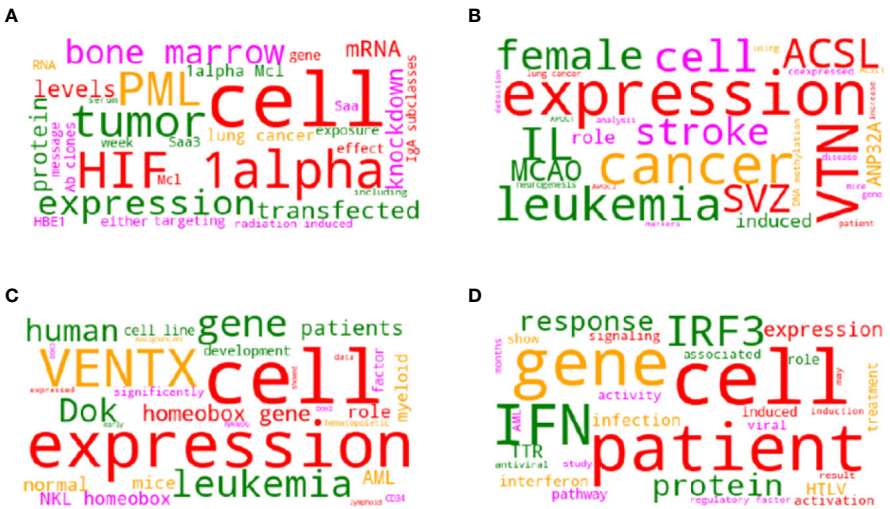


FIGURE 8 | The group of proteins present in each time point and at different levels of expression, return frequencies of different terms in Pubmed search. Word cloud generated by the frequency of the terms present in the abstracts returned as a result of the search in the Pubmed. Four searches were carried out with the term leukemia and proteins in each analyzed condition [**(A)** - D0 down, **(B)** - D28 down, **(C)** - D0 up, and **(D)** - D28 up], returning the 10 most recent works. The abstracts of the works were concatenated and organized as a term corpus, analyzed by the Natural Language Toolkit package - NLTK 3.4 and WordCloud 1.6.0 for Python. The size of the word represents the frequency in the corpus.

patients should be important for their remission since this protein is connected with angiogenesis, tumor growth, and metastasis (28).

Another important chemotherapy-induced event reported here refers to the super expression of the transcription factor IRF3 in the blood of B-ALL-patients, induced by chemotherapy. From the best of our knowledge, this is the first report that describes the presence of IRF3 in the blood of B-ALL patients. Despite the role of IRF3 in B-ALL is little known, it seems to be linked to B-cell differentiation in the presence of the TEL-AML1 fusion protein, the most common genetic rearrangement reported in B-ALL patients (29). Considering this, IRF3 could represent a positive mechanism induced by chemotherapy in the present study, since after killing the malignant clones of B-ALL, the main systemic signaling changed to cell differentiation. To understand the impact of chemotherapy on IRF3 activation and B-ALL dynamics during induction, we measured the levels of IFN-γ, a type I interferon, that can both regulate (30) and be

regulated by IRF3 (31). The axis IRF3/IFN-γ has been described as a pivotal immune mechanism in anti-tumor responses (31). Our data have shown that, at the end of the treatment, there was a considerable reduction in the PB/BM ratio of IFN-γ. In this context, the possibilities are i) the IFN-γ consumption by immune cells against B-ALL, since at D28 the tumor cells were cleared from BM by chemotherapy (as shown in **Table 1**, mean tumor cells counting in bone marrow aspirate under 7%) and ii) that IRF3 super expression at D28 was enough to induce the production of IFN-γ by cells from BM niche, more than those in the PB (as demonstrated in the box of **Figure 7**).

This reasoning takes our discussion to another super expressed protein found in the blood of our patients, the transthyretin. This protein has a role in normal cell differentiation by providing adequate levels of thyroxin, a hormone requested for B-cell maturation (32). Transthyretin has been reported as low at B-ALL diagnosis, but progressively augments in ALL patients during the induction phase of treatment (33), suggesting that this protein can be induced by chemotherapy. Investigation of transthyretin levels at D28 shown that its PB/BM ratio is significantly higher than at D0, corroborating this hypothesis. Considering this, it could be suggested that transthyretin super expression after induction treatment could represent a positive factor associated with B-ALL eradication. High transthyretin observed at D28 was concomitant with the elimination of the immature malignant clones in both PB and BM in all patients, suggesting that its augment could represent an additional marker for tumor cell elimination during the B-ALL induction phase treatment.

Therefore, D28 proteomic profiling called attention to proteins that are enrolled in the resolution of B-ALL by regulating cell proliferation and differentiation homeostasis.

The *in silico* analysis evidenced the possible TFs related to disease (D0) and to the induction treatment (D28). This because (1) most of the shared TFs among the different endpoints (D0 versus D28) do not have the same representativeness, suggesting that their regulation is altered with treatment. As shown from our data, AML1 (RUNX1) exhibited different targets and crosstalk to other TFs, despite being a TF present in both D0 and D28 analysis. This TF is mostly in AML due to the ETV6/RUNX1 translocation and its relevance has been recently discussed in B-ALL (34, 35). However, the RUNX1 role is still not investigated for non-translocated tumors. Secondly, the "exclusive" TFs from each data may account for the more specific molecular changes in B-ALL before and after treatment. TAL1 (TAL BHLH Transcription Factor 1, Erythroid Differentiation Factor) has been described classically as associated with ALL-T and enrolled with MAPK-Erk Pathway (36) and NF-kappaB Signaling (37). However, it has been reported as methylated in specific cohorts (patients aged >9 years and in patients showing relapse), suggesting its potential prognostic value (38). Our proteomic data together with bioinformatics analysis corroborates this hypothesis.

Proteomics has been used as a valuable tool to understand B-ALL biology and propose putative new markers of clinical relevance. A study from Cavalcante et al. conducted in B-ALL patients after induction chemotherapy using mass spectrometry reported proteins shared to those that we highlighted in the present study, including components from the coagulation pathway, pointed out as candidates to follow up favorable responses after induction therapy (39). Moreover, substantial advances are being reached with this approach regarding determinant clinical features of B-ALL, as the relation between surface proteins expression and disease risk stratification (40), highlighting proteomics as a complementary tool that can help to guide decision making.

In conclusion, our proteomic study added to understanding B-ALL biology at diagnosis and highlighted some important proteins and processes that may contribute to our understanding of the mechanisms concerning the impact of chemotherapy on disease resolution. Our findings highlight the relevance of IRF3-IFN γ axis induction as a possible mechanism enrolled in disease resolution, and point out transthyretin as an upregulated protein induced by the induction phase of chemotherapy.

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: The full list of identified proteins, together with raw data and details regarding methods were deposited in ProteomeXChange repository (Project accession: PXD021584, Project DOI: 10.6019/PXD021584).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by State University of Londrina/National Concil of Ethics - Conep. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CP and EA–conceptualization, methodology and investigation, data curation, and writing. FT and GB–conceptualization, patient selection, investigation, and writing. SC and GS–formal analysis and writing. All authors contributed to the article and approved the submitted version.

FUNDING

Araucária Foundation, Grant No 026/2016-PRPPG; Programa de Pesquisa para o SUS – PPSUS, Grant 01/2016; Conselho

Nacional de Desenvolvimento Tecnoloógico – CNPq, Grant Universal 14/2014; Coordenação de Aperfeiçoamento Pessoal de Nível Superior – CAPES.

ACKNOWLEDGMENTS

The authors are grateful to Conselho Nacional de Desenvolvimento Tecnológico (CNPq), INCT para o Controle do Câncer and Fundação Araucária for providing the financial support of the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.550213/full#supplementary-material

REFERENCES

1. Iacobucci I, Mullighan CG. Genetic Basis of Acute Lymphoblastic Leukemia. *J Clin Oncol* (2017) 35(9):975–83. doi: 10.1200/JCO.2016.70.7836
2. Singh N, Frey NV, Grupp SA, Maude SL. CAR T Cell Therapy in Acute Lymphoblastic Leukemia and Potential for Chronic Lymphocytic Leukemia. *Curr Treat Options Oncol* (2016) 17(6):28. doi: 10.1007/s11864-016-0406-4
3. Batista CR, Lim M, Laramée AS, Abu-Sardanah F, Xu LS, Hossain R, et al. Driver mutations in Janus kinases in a mouse model of B-cell leukemia induced by deletion of PU.1 and Spi-B. *Blood Adv* (2018) 2(21):2798–810. doi: 10.1182/bloodadvances.2018019950
4. Polak R, de Rooij B, Pieters R, den Boer ML. B-cell precursor acute lymphoblastic leukemia cells use tunneling nanotubes to orchestrate their microenvironment. *Blood* (2015) 126(21):2404–14. doi: 10.1182/blood-2015-03-634238
5. Luna-Aguirre CM, de la Luz Martinez-Fierro M, Mar-Aguilar F, Garza-Veloz I, Treviño-Alvarado V, Rojas-Martinez A, et al. Circulating microRNA expression profile in B-cell acute lymphoblastic leukemia. *Cancer Biomark* (2015) 15(3):299–310. doi: 10.3233/CBM-150465
6. Koskenvuo M, Ekman I, Saha E, Salokannel E, Matomäki J, Ilonen J, et al. Immunological Reconstitution in Children After Completing Conventional Chemotherapy of Acute Lymphoblastic Leukemia is Marked by Impaired B-cell Compartment. *Pediatr Blood Cancer* (2016) 63(9):1653–6. doi: 10.1002/pbc.26047
7. Garza-Veloz I, Martinez-Fierro ML, Jaime-Perez JC, Carrillo-Sanchez K, Ramos-Del Hoyo MG, Lugo-Trampe A, et al. Identification of differentially expressed genes associated with prognosis of B acute lymphoblastic leukemia. *Dis Markers* (2015) 2015:828145. doi: 10.1155/2015/828145
8. Yu R, Zhang J, Zang Y, Zeng L, Zuo W, Bai Y, et al. iTRAQ-based quantitative protein expression profiling of biomarkers in childhood B-cell and T-cell acute lymphoblastic leukemia. *Cancer Manag Res* (2019) 11:7047–63. doi: 10.2147/CMAR.S210093
9. Schroeder MP, Bastian L, Eckert C, Gökbuget N, James AR, Tanchez JO, et al. Integrated analysis of relapsed B-cell precursor Acute Lymphoblastic Leukemia identifies subtype-specific cytokine and metabolic signatures. *Sci Rep* (2019) 9(1):4188. doi: 10.1038/s41598-019-40786-1
10. Dehghan-Nayeri N, Eshghi P, Pour KG, Rezaei-Tavirani M, Omrani MD, Gharehbaghian A. Differential expression pattern of protein markers for predicting chemosensitivity of dexamethasone-based chemotherapy of B cell acute lymphoblastic leukemia. *Cancer Chemother Pharmacol* (2017) 80 (1):177–85. doi: 10.1007/s00280-017-3347-0
11. Calderon-Rodríguez SI, Sanabria-Salas MC, Umaña-Perez A. A comparative proteomic study of plasma in Colombian childhood acute lymphoblastic leukemia. *PloS One* (2019) 14(8):e0221509. doi: 10.1371/journal.pone.0221509
12. Brandalise SR, Pinheiro VR, Lee MLM. *Protocolo de Tratamento da Leucemia Linfoblástica Aguda da Criança e do Adolescente GBTLI LLA-2009*. São Paulo: Campinas: Grupo Sociedade Brasileira de Oncologia pediátrica (2011). p. 1–347.
13. Corrêa S, Panis C, Binato R, Herrera AC, Pizzatti L, Abdelhay E. Identifying potential markers in Breast Cancer subtypes using plasma label-free proteomics. *Proteomics* (2017) 151:33. doi: 10.1016/j.jprot.2016.07.030
14. Liu H, Finch JW, Luongo JA, Li GZ, Gebler JC. Development of an online two-dimensional nano-scale liquid chromatography / mass spectrometry method for improved chromatography performance and hydrophobic peptide recovery. *J Chromatogr* (2006) 1135:43–51. doi: 10.1016/j.chroma.2006.09.030
15. Pizzatti L, Panis C, Lemos G, Rocha M, Cecchini R, Souza HG, et al. Label-free MSE Proteomic Analysis of Chronic Myeloid Leukemia Bone Marrow Plasma: disclosing New Insights From Therapy Resistance. *Proteomics* (2012) 12:2618. doi: 10.1002/pmic.201200066
16. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, et al. Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem* (2005) 77:2187. doi: 10.1021/ac048455k
17. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics* (2006) 5:144. doi: 10.1074/mcp.M500230-MCP200
18. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* (2015) 43:D447. doi: 10.1093/nar/gku1003
19. Panis C, Pizzatti L, Abdelhay E. How can Proteomics Reach Cancer Biomarkers? *Curr Proteom.* (2013) 10:136–49. doi: 10.2174/1570164611310020006
20. Panis C, Pizzatti L, Souza GF, Abdelhay E. Clinical proteomics in cancer: Where we are. *Cancer Lett* (2016) 382(2):231–9. doi: 10.1016/j.canlet.2016.08.014
21. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol* (2014) 15(12):786–801. doi: 10.1038/nrm3904
22. Cameron W, Mojares E, Hernandes A. Role of Extracellular matrix in Development and Cancer Progression. *Int J Mol Sci* (2018) 19(10):3028. doi: 10.3390/ijms19103028
23. Mark D, Werb Z. How matrix metalloproteinases regulate cell behavior. *Annu Rev Cell Dev Biol* (2001) 17:463–516. doi: 10.1146/annurev.cellbio.17.1.463
24. Bazzi Z, Lanoue D, El-Youssef M, Romagnulo R, Tubman J, Cavallo-Medved D, et al. Activated thrombin-activatable fibrinolysis inhibitor (TAFIa) attenuates breast cancer cell metastatic behaviors through inhibition of plasminogen activation and extracellular proteolysis. *BMC Cancer* (2016) 16:328. doi: 10.1186/s12885-016-2359-1
25. Malik R, Lelkes PI, Cukierman E. Biomechanical and biochemical remodeling of stromal extracellular matrix in cancer. *Trends Biotechnol* (2015) 33(4):230–6. doi: 10.1016/j.tibtech.2015.01.004
26. Hurt EM, Chan K, Serrat MAD, Thomas SB, Veenstra TD, Farrar WL. Identification of Vitronectin as an Extrinsic Inducer of Cancer Stem Cell Differentiation and Tumor Formation. *Stem Cells* (2010) 28(3):390–8. doi: 10.1002/stem.27
27. Panis C, Pizzatti L, Bufalo AC, Herrera AC, Victorino VJ, Cecchini R, et al. Early downregulation of acute phase proteins after doxorubicin exposition in patients with breast cancer. *Tumour Biol* (2016) 37(3):3775–83. doi: 10.1007/s13277-015-4203-7
28. Tsopanoglou NE, Maragoudakis ME. Role of thrombin in angiogenesis and tumor progression. *Semin Thromb Hemost.* (2004) 30(1):63–9. doi: 10.1055/s-2004-822971
29. Laurentiis A, Hiscott J, Alcalay M. The TEL-AML1 fusion protein of acute lymphoblastic leukemia modulates IRF3 activity during early B-cell differentiation. *Oncogene* (2015) 34(49):6018–28. doi: 10.1038/onc.2015.50
30. Guinn ZP, Petro TM. Interferon regulatory factor 3 plays a role in macrophage responses to interferon-γ. *Immunobiology* (2019) 224(4):565–74. doi: 10.1016/j.imbio.2019.04.004
31. Guinn Z, Brown DM, Petro TM. Activation of IRF3 contributes to IFN-γ and ISG54 expression during the immune responses to B16F10 tumor growth. *Int Immunopharmacol* (2017) 50:121–9. doi: 10.1016/j.intimp.2017.06.016

32. Bloise FF, Oliveira FL, Nobrega AF, Vasconcellos R, Cordeiro A, Paiva LS, et al. High levels of circulating triiodothyronine induce plasma cell differentiation. *J Endocrinol* (2014) 220(3):305–17. doi: 10.1530/JOE-13-0315
33. Koskelo EK, Saarinen UM, Siimes MA. Low levels of serum transport proteins indicate catabolic protein status during induction therapy for acute lymphoblastic leukemia. *Pediatr Hematol Oncol* (1991) 8(1):53–9. doi: 10.3109/08880019109033427
34. Vijayakrishnan J, Qian M, Studd JB, Yang W, Kinnersley B, Law PJ, et al. Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. *Nat Commun* (2019) 10(1):5348. doi: 10.1038/s41467-019-13069-6
35. Chatterton Z, Morenos L, Mechinaud F, Ashley DM, Craig JM, Sexton-Oates A, et al. Epigenetic deregulation in pediatric acute lymphoblastic leukemia. *Epigenetics* (2014) 9(3):459–67. doi: 10.4161/epi.27585
36. Panepucci RA, Oliveira LH, Zanette DL, Viu Carrara Rde C, Araujo AG, Orellana MD, et al. Increased levels of NOTCH1, NF-kappaB, and other interconnected transcription factors characterize primitive sets of hematopoietic stem cells. *Stem Cells Dev* (2010) 19(3):321–32. doi: 10.1089/scd.2008.0397
37. Zhou RQ, Wu JH, Gong YP, Guo Y, Xing HY. Transcription factor SCL/TAL1 mediates the phosphorylation of MEK/ERK pathway in umbilical cord blood $CD34^+$ stem cells during hematopoietic differentiation. *Blood Cells Mol Dis* (2014). doi: 10.1016/j.bcmd.2013.12.007
38. Musialik E, Bujko M, Kober P, Wypych A, Gawle-Krawczyk K, Matysiak M, et al. Promoter methylation and expression levels of selected hematopoietic genes in pediatric B-cell acute lymphoblastic leukemia. *Blood Res* (2015) 50 (1):26–32. doi: 10.5045/br.2015.50.1.26
39. Cavalcante Mde S, Torres-Romero JC, Lobo MDP, Moreno FBMB, Bezerra LP, Lima DS, et al. A panel of glycoproteins as candidate biomarkers for early diagnosis and treatment evaluation of B-cell acute lymphoblastic leukemia. *Biomark Res* (2016) 4:1. doi: 10.1186/s40364-016-0055-6
40. Mirkowska P, Hofmann A, Sedek L, Slamova L, Mejstrikova E, Szczepanski T, et al. Leukemia surfaceome analysis reveals new disease-associated features. *Blood* (2013) 121(25):e149–59. doi: 10.1182/blood-2012-11-468702

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 © 2020 Broto, Corrêa, Trigo, Santos, Tomiotto-Pelissier, Pavanelli, Silveira, Abdelhay and Panis. 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 Oncology

ORIGINAL RESEARCH
published: 15 December 2020
doi: 10.3389/fonc.2020.523989

Proteomic Analysis of Pre-Invasive Serous Lesions of the Endometrium and Fallopian Tube Reveals Their Metastatic Potential

Mitchell Acland[1], Georgia Arentz[1], Max Mussared[2], Fergus Whitehead[3], Peter Hoffmann[4], Manuela Klingler-Hoffmann[4]* and Martin K. Oehler[4,5,6]**

[1] Adelaide Proteomics Centre, School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia, [2] School of Mathematical Sciences, The University of Adelaide, Adelaide, SA, Australia, [3] Cytopathology Department, Clinpath Pathology, Adelaide, SA, Australia, [4] Future Industries Institute, Mawson Lakes Campus, University of South Australia, Adelaide, SA, Australia, [5] Department of Gynaecological Oncology, Royal Adelaide Hospital, North Terrace, Adelaide, SA, Australia, [6] Robinson Research Institute, Discipline of Obstetrics and Gynaecology, Adelaide Medical School, The University of Adelaide, SA, Australia

OPEN ACCESS

Edited by:
Harsha Gowda,
The University of Queensland,
Australia

Reviewed by:
Wa Xian,
University of Connecticut,
United States
Jian-Jun Wei,
Northwestern University,
United States

***Correspondence:**
Martin K. Oehler
martin.oehler@adelaide.edu.au
Peter Hoffmann
peter.hoffmann@unisa.edu.au
Manuela Klingler-Hoffmann
Manuela.Klingler-Hoffmann@unisa.edu.au

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 17 January 2020
Accepted: 28 October 2020
Published: 15 December 2020

Citation:
Acland M, Arentz G, Mussared M, Whitehead F, Hoffmann P, Klingler-Hoffmann M and Oehler MK (2020) Proteomic Analysis of Pre-Invasive Serous Lesions of the Endometrium and Fallopian Tube Reveals Their Metastatic Potential. Front. Oncol. 10:523989. doi: 10.3389/fonc.2020.523989

Serous endometrial cancer (SEC) and high grade serous ovarian cancer (HGSOC) are aggressive gynecological malignancies with high rates of metastasis and poor prognosis. Endometrial intraepithelial carcinoma (EIC), the precursor for SEC, and serous tubal intraepithelial carcinoma (STIC), believed to be the precursor lesion for HGSOC, can also be associated with intraabdominal spread. To provide insight into the etiology of these precancerous lesions and to explore the potential molecular mechanisms underlying their metastatic behavior, we performed a proteomic mass spectrometry analysis in a patient with synchronous EIC and STIC. Through histological and molecular identification of precancerous lesions followed by laser capture microdissection, we were able to identify over 450 proteins within the precancerous lesions and adjacent healthy tissue. The proteomic analysis of STIC and EIC showed remarkable overlap in the proteomic patterns, reflecting early neoplastic changes in proliferation, loss of polarity and attachment. Our proteomic analysis showed that both EIC and STIC, despite being regarded as premalignant lesions, have metastatic potential, which correlates with the common presentation of invasive serous gynecological malignancies at advanced stage.

Keywords: proteomics, serous tubal intraepithelial carcinoma, endometrial intraepithelial carcinoma, serous endometrial carcinoma, high grade serous ovarian carcinoma

INTRODUCTION

Endometrial cancer (EC) is the 6th most common cancer in women worldwide and is the most common gynecological malignancy (1). Despite significant advances in early detection (2), molecular subtyping (3, 4), and new or improved treatment regimens (5), the relative survival of patients with EC has declined in recent times (6, 7).

Serous endometrial carcinoma (SEC) is a highly aggressive malignancy (8). It represents only 10% of EC cases (9, 10) but is responsible for 39% of EC related deaths (11) and is frequently diagnosed at late stage when prognosis is poor (9, 10). The current model of SEC development suggests that it evolves from a pre-neoplastic lesion in atrophic endometrium called endometrial glandular dysplasia (EmGD) (12, 13). This lesion then progresses further into endometrial intraepithelial carcinoma (EIC) and, finally, into SEC (14).

EmGD is characterized by loss of cell polarity, nuclear atypia, and nuclear hyperchromasia (12). It has a marked loss of heterozygosity in TP53 and chromosome p1; however, this is to a lesser extent than that seen in EIC and SEC (12). Although quite difficult to identify, nucleomegaly and staining for p53, MIB-1 as well as IMP3 are characteristics of EmGD (14). The connection between EmGD and SEC has been confirmed through identical mutations observed in EmGD and subsequent SEC (15).

EIC was first described in the 1990s (16) and seen to arise almost exclusively in atrophic endometrium and in the context of SEC in a majority of cases (17). EIC exhibit similar features as EmGD but with further nucleomegaly, nuclear irregularity and hyperchromasia (14). However, cases of EIC associated with extrauterine metastasis suggest that EIC may more closely resemble SEC (18, 19). This early peritoneal spread is in stark contrast to non-serous endometrial cancers which usually do not show early peritoneal spread but preferentially invade the myometrium and spread to the lymph nodes (18, 19).

SEC closely resembles other serous cancers of the female genital tract, such as high grade serous ovarian cancer (HGSOC) (4). Both diseases share similar molecular features and clinical properties (20) and, consequently, are treated in a similar way (21, 22).

In recent years the fallopian tube has been identified as the precursor site of HGSOC, specifically serous tubal intraepithelial carcinoma (STIC) (23–27). These lesions are identified histopathologically by a "p53 signature" comprised of strong p53 staining (28), p53 mutations (29), positive γ-H2AX staining (indicating DNA damage) and lack of Ki-67 staining (indicating low proliferation) (30) (**Table 1**). STICs share many genomic features with HGSOC, such as genomic instability (31, 32), and HGSOC has a gene expression profile more similar to the fallopian epithelium than the ovarian surface epithelium (33).

Results from studies using mouse models have established a connection between STIC and subsequent HGSOC (34, 35). Inactivation of PTEN, p53 and BRCA1/2 in the fallopian tubes of mice resulted in STIC and concurrent HGSOC with ovarian and peritoneal spread (34, 35). However, in the absence of BRCA1/2 inactivation, STIC developed but did not progress to metastatic HGSOC in this mouse model (34).

Cases of HGSOC arising in the absence of STIC (35–37) have also been reported suggesting that other, as yet unidentified, precursor lesions might exist (38). "Early Serous Tubal Proliferations" (ESTP) have been identified as a potential HGSOC precursor. They are found in the fimbria (30), demonstrate DNA damage (30), and are found in non-ciliated cells, which also give rise to STIC (39). A physical and lineage continuity has been demonstrated between ESTP and STIC suggesting that some ESTP give rise to STIC and subsequently to HGSOC (30, 32).

The understanding that HGSOC arises from the fallopian tube in many cases has changed the understanding of serous ovarian cancer. Now serous cancers of the fallopian tube, peritoneum and ovary are thought to share a common origin in the fallopian tube (20). While it is well established that many HGSOC arise from the fallopian tube, it has not been excluded that some serous cancers of the endometrium and ovary may share common origins. For example, Roelofsen et al. (40) suggested that some serous ovarian cancers (SOC) may arise from EIC by showing that they shared TP53 mutations, similar expression of p53, Ki67, estrogen, and progesterone receptors (40). Additionally, Tolcher et al. (20) analyzed 38 patients with SEC and investigated their fallopian tubes. They found STIC, without evidence of tubal metastasis, in 2 of these cases (20).

To better understand the potential link between serous preinvasive lesions of the female genital tract and serous gynaecological cancer, molecular investigations of the precursor lesions of the endometrium and fallopian tube are required. Here, we present the first proteomic analysis of synchronous precancerous lesions of the endometrium and fallopian tubes in a patient without invasive malignancy, by means of mass spectrometry. This precludes the possibility of these premalignant lesions representing metastases from established primary tumors. The analysis of EIC and STIC in this context provides insight into the temporal and mechanistic features of their development and dissemination.

MATERIALS AND METHODS

Sample

Archived formalin-fixed paraffin-embedded (FFPE) fallopian tube and endometrial tissues from a 67-year-old female who

TABLE 1 | Morphological and molecular features of precancerous lesions of the gynecological tract.

Location	Precancerous lesion	Morphological features	Molecular features
Endometrium	EmGD	Some nuclear atypia, some nucleomegaly.	Low rates of loss of heterozygosity in TP53 and chromosome P1. p53 mutations in 50% of cells. Staining for P53, MIB1, and IMP3.
	EIC	Extensive nuclear atypia, extensive nucleomegaly, hyperchromasia.	High rates of loss of heterozygosity in TP53 and chromosome P1. p53 mutations in 75% of cells. Strong staining for P53, MIB1, and IMP3.
Fallopian tube	ESTP	No visible morphological features.	TP53 mutations, p53 staining, DNA damage.
	STIC	Hyperchromasia, nucleomegaly.	TP53 mutations, p53 and MIB1 staining, DNA damage, Chromosomal instability.

had undergone a total abdominal hysterectomy and bilateral salpingo-oophorectomy for endometrial hyperplasia was retrieved for analysis with approval of the Research Ethics Committee of the Royal Adelaide Hospital. The fallopian tube and endometrial tissues were processed using standard procedures, stained with hematoxylin and eosin, and annotated by a pathologist. P53 and MIB1 immunostaining was also performed to confirm the location of precancerous lesions.

Laser Microdissection and Sample Preparation

FFPE tissues were sectioned at 4-µm thickness, water bath mounted onto PEN membrane slides (Micro-Dissect, Herborn, Germany), and deparaffinized by submersion in xylene for 5 min, following by two 2-min incubations in 100% ethanol, and two 5-min incubations in water. Areas of STIC, EIC and adjacent healthy epithelium were dissected using a Leica AS LCM microscope (Leica Microsystems, Wetzlar, Germany) into 20 µl of 10 mM citric acid buffer (pH = 6) and subjected to heat induced antigen retrieval by incubation at 100°C for 90 min. The solution containing the protein extracts were digested with trypsin gold (Promega, Madison, WI, USA) as described in Mittal et al. (41) using a modified FASP method (42). In brief, protein extracts were mixed with 0.2 ml of 8M urea in 0.1M Tris/HCl, pH 8.5 before being loaded into a 30k Microcon filtration device (Millipore) and centrifuged at 14,000g for 15 min. This step was repeated to ensure the removal of any residual contaminants. Samples were reduced with 5 mM DTT (Roche) for 45 min at room temperature and alkylated with 10mM iodoacetamide (IAA) (GE Healthcare, Little Chalfont, UK) for 30 min at room temperature in the dark followed by centrifugation at 14,000g for 15 min. The protein concentrate was diluted with 0.2mL of 8M urea in 0.1M Tris/HCl, pH 8.5, and spun again at 14,000g for 15 min. This step was repeated twice. Samples were buffered with 10mM NH_4HCO_3 and digested with 100ng trypsin gold overnight at 37°C. Peptides were collected by centrifugation of the filter unit at 14,000g for 20 min.

Nanoflow Liquid Chromatography Tandem Mass Spectrometry

Nanoflow liquid chromatography tandem mass spectrometry (Nano-LC-MS/MS) was performed on each sample in duplicate using an Ultimate 3000 RSLC system (Thermo-Fisher Scientific, Waltham, USA) coupled to an Impact HD™ QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany) *via* an Advance CaptiveSpray source (Bruker Daltonics). Peptide samples were pre-concentrated onto a C18 trapping column (Acclaim PepMap100 C18 75 µm × 20 mm, Thermo-Fisher Scientific) at a flow rate of 5 µl/min in 2% (v/v) ACN 0.1% (v/v) TFA for 10 min. Peptide separation was performed using a 75 µm ID C18 column (Acclaim PepMap100 C18 75 µm × 50 cm, Thermo-Fisher Scientific) at a flow rate of 0.2 µl/min using a linear gradient from 5 to 45% B (A: 5% (v/v) ACN 0.1% (v/v) FA, B: 80% (v/v) ACN 0.1% (v/v) FA) over 130 min, followed by a 20-min wash with 90% B, and a 20-min equilibration with 5% A. MS scans were acquired in the mass range of 300 to 2,200 m/z in a data-dependent fashion using Bruker's Shotgun Instant Expertise™ method. This method uses IDAS (intensity dependent acquisition speed) to adapt the speed of acquisition depending on the intensity of precursor ions (fixed cycle time), and RT^2 (RealTime Re-Think) to exclude previously selected precursor ions from undergoing re-fragmentation unless the chromatographic peak intensity of the ion has increased by a factor of 5. Singly charged precursor ions were excluded from acquisition. Collision energy ranged from 23% to 65% as determined by the m/z of the precursor ion.

Data Analysis

Spectra were analyzed using the MaxQuant software (version 1.5.2.8) with the Andromeda search engine (43) against the UniProt non-redundant human database. The standard Bruker QTOF settings in MaxQuant were used with a mass error tolerance of 40 ppm. The variable modifications of oxidation of methionine and the fixed modification of carbamidomethyl of cysteines were specified, with the digestion enzyme specified as trypsin. The protein false discovery rate (FDR) and peptide spectrum match FDRs were both set to 1% using a target decoy approach, with a minimum peptide length of 7 amino acids (43). Only unique and razor peptides were used when reporting protein identifications.

Gene Expression Analysis

In order to assess the gene expression levels of the corresponding proteins of interest in early stage I ovarian carcinoma tissues, the dataset of Yoshihara et al. (44) [Gene Expression Omnibus (GEO) Accession GSE12470, http://www.ncbi.nlm.nih.gov/geo/] was considered. From this dataset the expression of EPCAM and CAPS was considered in 8 in early stage I patients compared to 10 healthy peritoneum control tissues. The results were natural log transformed and compared using paired T-tests and p-values < 0.05 were considered significant. Full patient details are available in the Yoshihara et al. (44) manuscript.

To assess the gene expression levels of the corresponding proteins of interest in differing subtypes of EOC, the dataset of Hendrix et al. (45) (GEO Accession GSE6008, http://www.ncbi.nlm.nih.gov/geo/) was considered. From this dataset the expression of TPPP3, SORD and VCAN was investigated in 37 endometrioid, 41 serous, 13 mucinous, and 8 clear cell ovarian carcinoma tissues, and 4 normal control tissues. Transformation of the data and full patient details are available from the Hendrix et al. (45) manuscript. Groups were compared using paired T-tests and p-values < 0.05 were considered significant.

With the aim to assess the gene expression levels of corresponding proteins in early stage 1 endometrial carcinoma tissue, R was used to investigate the data set of Days et al. (46) (GEO Accession GDS4589, http://www.ncbi.nlm.nih.gov/geo/). This this dataset the expression of EPCAM and CAPS were investigated in 79 endometrioid and 12 serous papillary endometrial carcinoma tissues, and in 12 normal control tissues. The results were natural log transformed and compared using paired T-tests. P-values < 0.05 were considered significant. Full patient details are available through the Day's et al. (46) manuscript.

To evaluate the gene expression levels of corresponding proteins in different subtypes of endometrial carcinoma tissue, R was used to investigate the data set of Kandolth et al. (4) (https://gdc.cancer.gov/node/875). From this data set the expression of TPPP3, SPATA18, ERO1A, SORD and VCAN were investigated in 13 CN high, 15 CN low, 16 MSI hypermutated, and 4 POLE ultra-mutated carcinoma tissues. Transformation of the data set is detailed in Kandolth et al. (4). Groups were compared using paired T-tests and p-values < 0.05 were considered significant. Full patient details are available in the Kandolth et al. (4) manuscript.

RESULTS

Histological Analysis Identifying STIC and EIC

Upon analysis of the H&E stained tissue sections, atypical intraepithelial proliferation involving a small population of cells in the endometrium were identified. These changes were consistent with early stages of EIC (**Table 1**). Similarly, atypical changes involving a small population of cells in the fimbriated tube (**Table 1**), consistent with early stages of STIC, were detected and both EIC and STIC are represented in **Figure 1**. Immunoperoxidase staining for p53 and MIB1 revealed atypical intraepithelial epithelial proliferations in both the fimbriated tube and endometrial lining, confirming the presence of STIC and EIC (data not shown).

Proteomic Comparison of Healthy Epithelia to STIC and EIC

Regions of STIC, EIC, and adjacent healthy epithelium were laser microdissected (LMD) from sectioned tubal and endometrial specimens and analyzed by Nano-LC-MS/MS. In total, 453 proteins were detected across the 4 tissue types (369 proteins in the STIC, 110 proteins in healthy tubal epithelium, 428 proteins in EIC, and 162 proteins in healthy endometrial epithelium (**Supplementary Table 1**).

When comparing the numbers of identified proteins, the greatest overlap occurred between STIC and EIC with 348 identical proteins identified in both samples. Across all tissue types 73 proteins were detectable, with 96 identical proteins identified in both the STIC and healthy tubal tissue, 157 identical proteins detected in both the EIC and healthy endometrial tissue, and 85 identical proteins were detected in both healthy tissues (**Figure 2**).

Proteins Relevant to STIC, EIC, and the Development of Gynecological Cancer

In analyzing potential links between STIC and EIC cells and their respective invasive carcinomas, two of the identified proteins were of particular interest based on their involvement in gynecological cancer development: Epithelial cell adhesion molecule (EPCAM) and Calcyphosin (CAPS) (47–49). EPCAM was identified in both the STIC and EIC tissue samples but was not detected in either of the healthy tissues while CAPS was only detected in the healthy tube and in STIC.

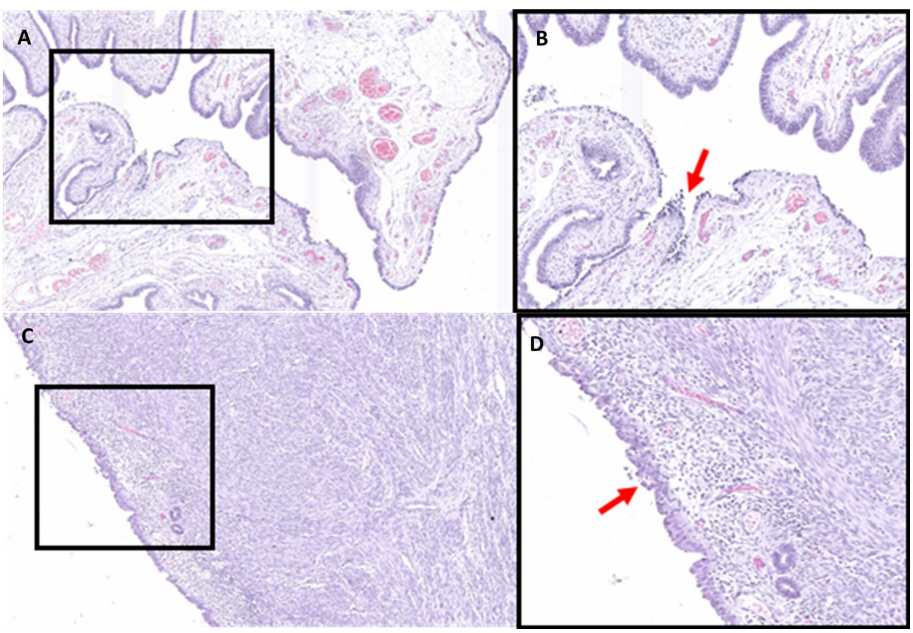


FIGURE 1 | Hematoxylin and Eosin stained fallopian tube **(A, B)** and endometrium tissue **(C, D)** at 6× **(A, C)** and 12× **(B, D)** magnification. Areas of STIC **(B)** and EIC **(D)** are indicated by the red arrows.

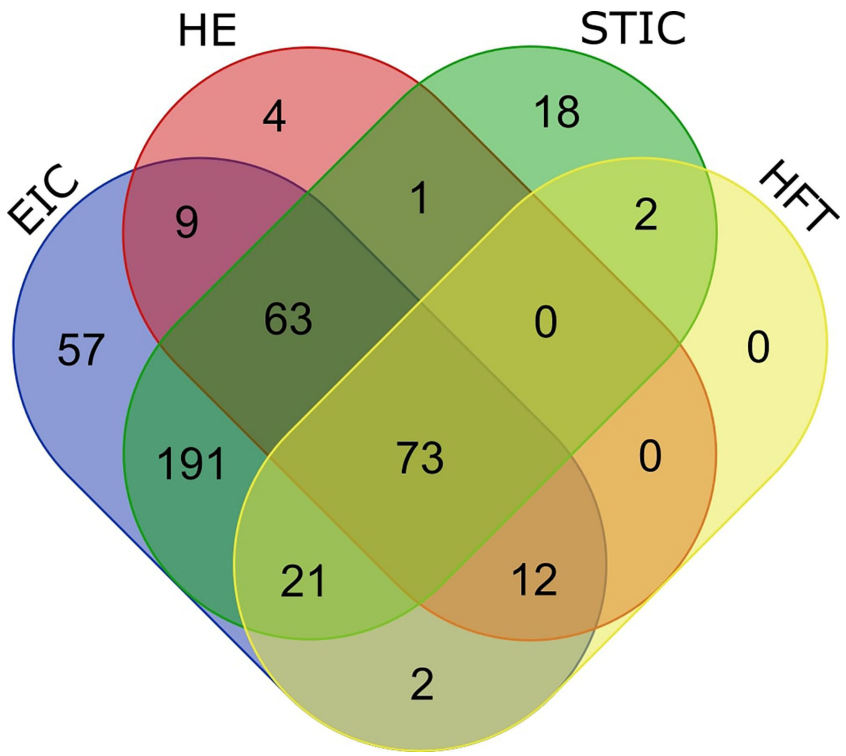


FIGURE 2 | Venn diagram describing overlapping protein identifications in endometrial intraepithelial carcinoma (EIC), healthy endometrium (HE), serous tubal intraepithelial carcinoma (STIC) and healthy fallopian tube (HFT). A significant overlap of (348 proteins) is observed between EIC and STIC (diagram generated at http://bioinformatics.psb.ugent.be/webtools/Venn/).

Expression of EPCAM and CAPS in Early Stage Serous Ovarian Carcinomas Compared to Healthy Peritoneal Tissue

To evaluate the expression of *EPCAM* and *CAPS* in early stage I serous ovarian carcinoma tissues, oligonucleotide microarray data was considered from the GEO data set GSE12470 (44). We chose to investigate the expression of these genes in early stage ovarian cancer as this is expected to be the stage following STIC in the development of serous ovarian cancer. STIC lesions were not deemed an appropriate comparison samples as they are often taken in the context of metastatic disease and potentially represent metastatic implants, and therefore more developed cancer, rather than true precursor lesions.

The gene expression levels were analyzed in 10 healthy peritoneal control tissues compared to 8 Stage I serous EOC tissues. The median expression levels of both *CAPS* and *EPCAM* were found to be significantly increased in the Stage I serous OC tissues compared to healthy peritoneal controls ($p = 0.00014$ and $p = 3.3 \times 10^{-7}$, respectively) (**Figure 3**).

Healthy OSE is infrequently available for research purposes given healthy ovaries are rarely removed during any type of medical procedure. Healthy peritoneum is an effective control because the lining of the ovaries is comprised of a single-cell mesothelial layer of poorly differentiated epithelium derived from the coelomic epithelium and extended to the serosa peritoneal cavity (50). In addition, it has been reported that peritoneal mesothelium and OSE are structurally very similar (51) and are both negative for *EPCAM* and *CAPS* expression (52, 53).

Expression of EPCAM and CAPS in Early Stage Endometrial Carcinoma Compared to Healthy Endometrial Tissue

To assess the expression of *EPCAM* and *CAPS* in early stage endometrial cancer, oligonucleotide microarray data was considered from the GEO data set GDS4589. The gene expression levels were analyzed from 79 endometrioid and 12 serous papillary endometrial carcinomas as well as 12 healthy controls. The median expression of CAPS was increased significantly in EEC ($p = 0.00066$), while it was significantly decreased in SEC ($p = 0.023$) compared to healthy endometrium. The median expression of EPCAM was significantly increased in EEC ($p = 0.02$) and further increased in SEC ($p = 0.0007$) compared to healthy controls (**Figure 4**).

Expression of Proteins Identified Exclusively in the STIC or EIC Across Ovarian Cancer Subtypes

Proteins detected exclusively in either the STIC or EIC tissues were analyzed to determine if their expression is specific to

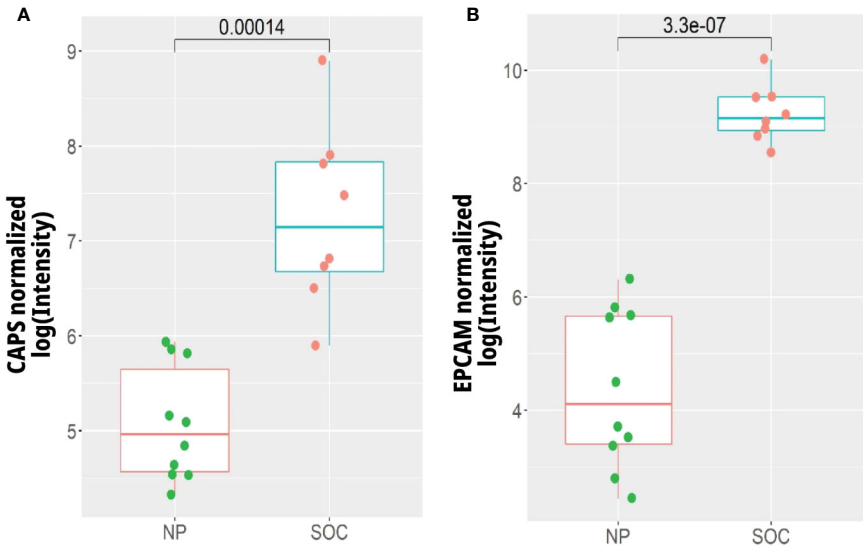


FIGURE 3 | Gene expression of **(A)** CAPS and **(B)** EPCAM in early stage I serous ovarian cancer tissues (SOC) (n = 8) compared to normal peritoneum (NP) (n = 10). Expression levels were extracted from the data of Yoshihara et al. (44) (GEO Accession GSE12470) *via* the R package CuratedOvarianData (http://www.ncbi.nlm.nih.gov/geo/).

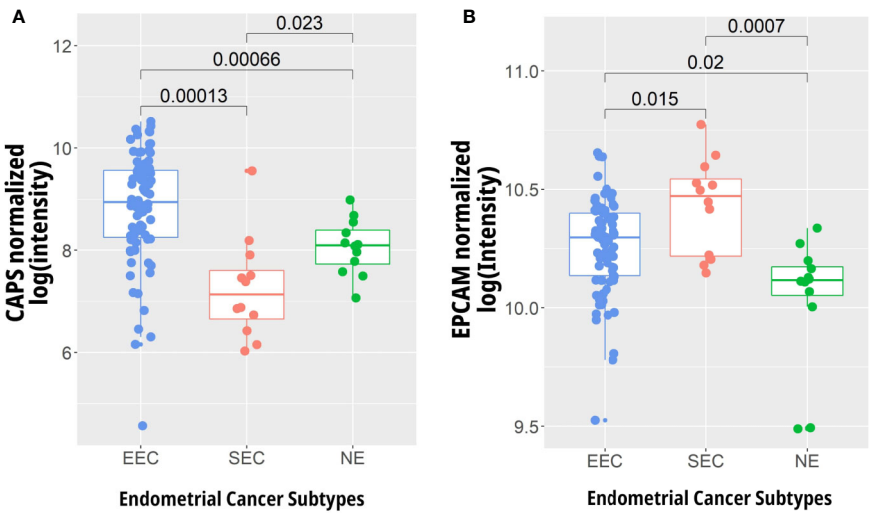


FIGURE 4 | Gene expression analysis of **(A)** CAPS and **(B)** EPCAM in early stage endometrioid (EEC) (n = 79) and serous (SEC) endometrial carcinoma (n = 12) compared to normal endometrium (NE) (n = 12). Expression levels were extracted from the data of Days et al. (46) (GEO Accession GDS4589, http://www.ncbi.nlm.nih.gov/geo/) using R.

certain gynecological tissues. Marker proteins expressed exclusively by the precancerous cells of either the tube or endometrium, which are also expressed by ovarian carcinomas, may aid in determining the tissue specific origin of HGSOC. Eighteen proteins were detected exclusively in the STIC tissue and 57 proteins in the EIC tissue. Of these proteins a small number appear to be specific to certain gynecological tissues according to Protein Atlas; 2 identified from STIC (MIEAP, TPPP3), and 3 identified from EIC (ERO1A, DHSO, CSPG2/*VCAN*). These proteins and their tissue specificities across all gynecological tissues are listed in **Table 2**. The remaining proteins appeared to be more homogenously expressed and hence were not analyzed further.

The gene expression levels corresponding to the proteins listed in **Table 2** were compared using oligonucleotide microarray data from the GEO data set GSE6008 (**Figure 5**).

TABLE 2 | Proteins identified exclusively in the STIC or EIC that appear to be specific to certain gynecological tissues.

Protein	Detected In	Protein abundance levels in healthy tissues *	Protein abundance in ovarian cancer *	Protein abundance in endometrial cancer *
Mitochondrial eating protein (MIEAP), *SPATA18*[#]	STIC	High in fallopian tube, low in endometrium and breast. Not expressed in the ovary.	Low	Low
Tubulin polymerization-promoting protein member 3 (TPPP3), *TPPP3*	STIC	Medium in glandular cells of the endometrium and cervix. Highly enriched in tube, negative in ovaries.	Medium	Medium to high
Endoplasmic reticulum oxidoreductin1-like protein alpha (ERO1A), *ERO1A*[#]	EIC	High in cervix and medium in endometrium. No expression in tube or ovary.	Low	Low
Sorbitol dehydrogenase (DHSO), *SORD*	EIC	High in endometrium, cervix/uterinus, and low in breast. Negative in tube and ovaries.	Low	Medium
Versican core protein (CSPG2), *VCAN*	EIC	High in placenta. Medium in cervix/uterus, and endometrium. Low in tube and ovaries.	Low to medium	Low to Medium

**According to Protein Atlas (http://www.proteinatlas.org/).*
[#]Not present in the Hendrix et al. data set. (45).

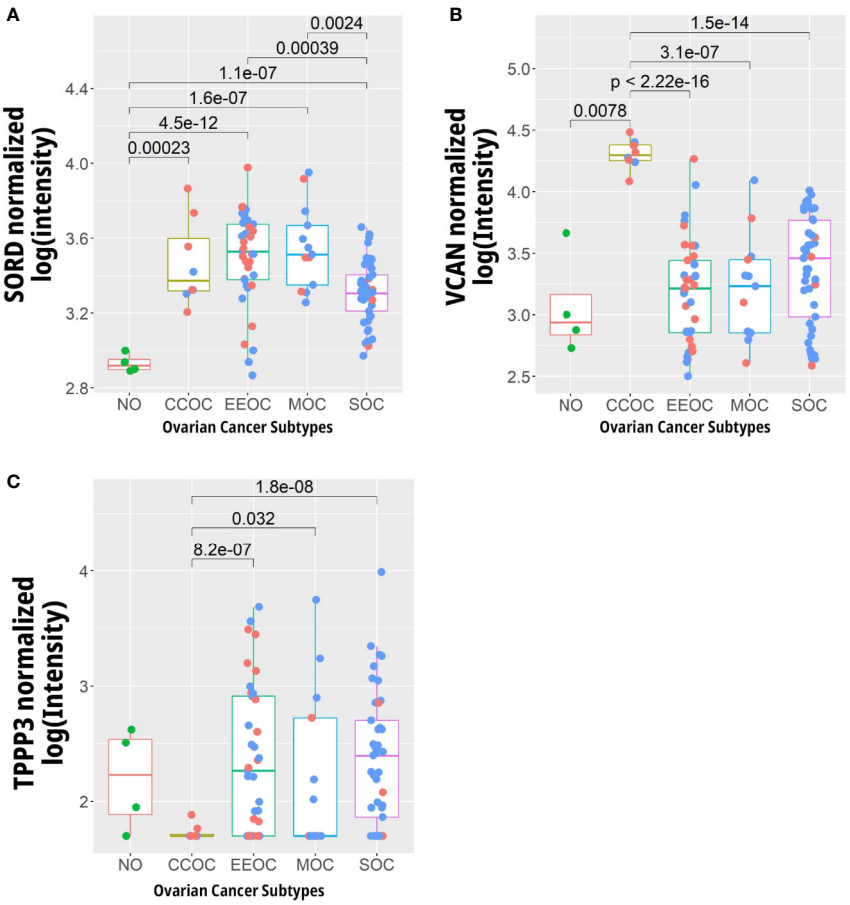


FIGURE 5 | Expression levels of genes whose protein abundance is specific to certain to gynaecological tissues. **(A)** SORD, **(B)** TPP3, and **(C)** VCAN in 4 normal ovarian (NO), 8 clear cell (CCOC), 37 endometrial (EEOC), 13 mucinous (MOC), and 41 serous (SOC) ovarian carcinomas. Blue dots represent data points from late stage patients while the red represent early stage. Data gathered from Hendrix et al. (45) (GEO Accession GSE6008, http://www.ncbi.nlm.nih.gov/geo/).

For the gene *TPPP3* there was reduced expression in ovarian clear cell carcinoma compared to healthy tissue and other ovarian cancer subtypes. *SORD* showed significantly lower expression levels in the normal control tissues, and to a smaller extent, in serous ovarian carcinoma compared to the other subtypes. The high expression of *SORD*, whose related protein was detected only in EIC in our data set, in non-SOC was unexpected as the protein atlas reports low expression of *SORD* in ovarian cancer. CSPG2/*VCAN* expression, which was detected at the protein level in EIC only, was increased in CCOC when

compared to the normal tissue and other EOC subtypes (**Figure 5**).

Expression of Proteins Identified Exclusively in STIC or EIC Across Endometrial Cancer Subtypes

By investigating the gene expression corresponding to proteins in our data set which are enriched in specific gynecological tissues, we aimed to investigate the connection between preneoplastic lesions of the endometrium and tube with subtypes of endometrial cancer as defined by the TCGA. The expression levels of the genes corresponding to the proteins listed in **Table 1** were compared using oligonucleotide microarray data from Kandolth et al. (4) (https://gdc.cancer.gov/node/875) (**Figure 6**).

They were selected as their related proteins were found exclusively in a precancerous lesion and their abundance

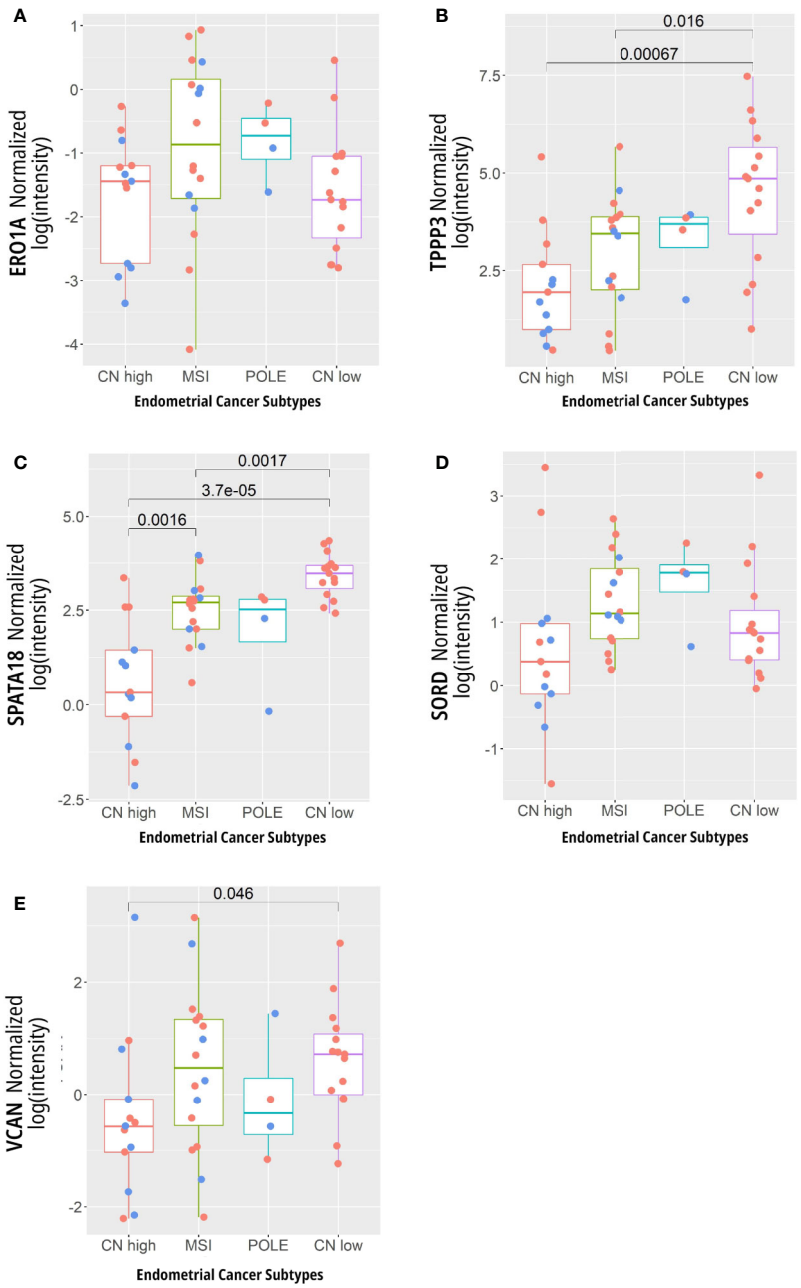


FIGURE 6 | Expression levels of genes whose proteins are associated with specific gynecological tissues. **(A)** ERO1A, **(B)** TPPP3, **(C)** SPATA18, **(D)** SORD, and **(E)** VCAN in 13 copy number (CN) high, 16 micro satellite instability (MSI) hypermutated, 4 POLE ultra-mutated, and 15 CN low. Here, the blue dots represent data points from late stage patients while the red represent early stage cancers. Data from Kandolth et al. (4) (https://gdc.cancer.gov/node/875).

(according to protein atlas) was unique to specific gynecological tissue. However, most did not show different gene expression between the TCGA defined subtypes, except for *TPPP3* and *SPATA18*, which both showed decreased expression in the CN high (serous) subgroup compared to other subtypes (**Figure 6**). These proteins were detected exclusively in STIC in our analysis and their expression was seen to be decreased in SEC compared to healthy controls (**Supplementary Figure 2**).

Proteins Associated With Metastasis Identified in STIC or EIC

The 34 identified proteins expressed exclusively in STIC and/or EIC which are known to be involved in metastasis or migration were listed in **Supplementary Table 2**. Most of these proteins are implicated in the promotion of metastasis, with the exceptions of Galectin-9 (54, 55), Mimecan (56, 57), and Catenin alpha-1 (58, 59) which have been implicated in the inhibition of metastasis.

DISCUSSION

The application of proteomic techniques, particularly that of mass spectrometry, hold the potential to provide a temporal snapshot of the molecular features within a given sample. Here, we provide what is, to the best of our knowledge, the first proteomic analyses of synchronous precursor lesions of serous endometrial and high-grade ovarian cancer. Through histological and molecular identification of precancerous lesions, followed by laser-capture microdissection and mass spectrometry analysis, we were able to identify over 450 proteins within the precancerous lesions and adjacent healthy tissue. The proteomic profiles of the precancerous lesions showed striking similarity (**Figure 1**) and shared a molecular profile indicative of metastatic transformation (**Supplementary Table 2**). To investigate the connection between these precancerous lesions and serous gynecological malignancies we investigated the gene expression of several proteins of interest in serous endometrial and ovarian cancer data sets (**Figures 3–6**).

EPCAM and CAPS were selected for further investigation based on their implication in the development of ovarian and endometrial cancer (47–49). Genomic analysis of their expression in Stage I SOC compared to peritoneal control tissue revealed significantly increased levels in the cancerous specimens (**Figure 3**). It is well recognized that EPCAM expression has a complex relationship to SOC development (60); however, it has not previously been identified in STIC or other gynecological precancerous lesions. This transmembrane adhesion molecule plays a role in migration and proliferation in wound healing (61) as well as the maintenance of pluripotency in stem cells (62).

The exact function of CAPS is unknown, but it is a suggested target of cAMP-dependent protein kinase and has been implicated in the cAMP and calcium-phosphatidylinositol signaling cascades (63). According to protein atlas, CAPS has high expression in the fallopian tube but is not expressed in healthy OSE (http://www.proteinatlas.org/ENSG00000152611-CAPSL/tissue) which is in agreement with our identification of this protein in both the healthy fallopian tissue and STIC. Interestingly, *CAPS* gene expression was seen to be increased in SOC compared to control tissue (**Figure 3**). Furthermore, *EPCAM* is not expressed in healthy ovarian surface epithelium (OSE) (53) but is frequently up regulated in ovarian cancer (60). The single case study presented here is insufficient to draw broader conclusions about potential markers of tissues of origin, but we believe that expression of EPCAM and CAPS merit further investigation in a larger cohort.

To further investigate proteins which may act as markers of tissue of origin in SEC or SOC, we investigated proteins identified exclusively in either the STIC or EIC which were unique to specific gynecological tissues according to protein atlas. CSGP/*VCAN* was identified exclusively in EIC, was seen to have increased expression in COCC compared to other ovarian cancer cell types (**Figure 5**) and had increased expression in SEC compared to healthy control tissue (**Supplementary Figure 2**). This protein has been previously seen to be increased in COCC (64, 65) and represents a potential link between EIC and COCC. COCC has been suggested to arise from endometriosis lesions (66) potentially representing a pathway for endometrial origin of COCC.

A previous mass spectrometry-based analysis of ovarian cancer precursor lesions was performed by Levenon et al. (67) and investigated *ex vivo* culture derived cells from fallopian tube fimbria (67). Their proteomic analysis identified 11 different ovarian cancer biomarkers present in this *ex vivo* model. A larger study by M. Eckart et al. (68) investigated both tumor and stroma tissue from STIC, invasive fallopian tube lesions, invasive ovarian lesions and omental metastasis (68). In addition to identifying N-methyltransferase (NNMT) as a metabolic regulator of cancer associated fibroblasts, they also showed that the molecular profiles of primary cancers and metastatic implants were remarkably similar within the same patient while the microenvironment showed site specific differences. The STIC lesion they investigated showed lower expression compared to normal tube in 4 of our proteins of interest (CAPS, ERO1A, TPPP3, and SPATA18) and similar expression in 2 (EPCAM and SORD). Only VCAN showed marginally increased expression in STIC compared to normal fallopian tube epithelium in this analysis. The authors' focus on stromal tissue, a lack of normal epithelial ovarian or omental controls and the potential that STIC represents metastatic lesions from a primary ovarian tumor limits further comparison to this data set.

The traditional understanding of cancer development is that it acquires metastatic capacities over time within the primary lesion (69). However, the identification of STIC as the origin site of HGSOC implicates that these pre-invasive lesions can leave their site of origin and establish themselves in distant locations. In addition, EIC is often identified with extrauterine spread (18, 19) which can take the form of EIC like growths on the ovaries, peritoneum and fallopian tube in the absence of obvious disease in these locations (16, 18). Together, this suggests that these premalignant lesions possess some migratory or metastatic ability facilitating their translocation to distant sites within the gynecological system. Here, we identified numerous metastasis and migration related proteins in precancerous lesions in the absence of malignant disease (**Supplementary Table 2**).

A major limitation of this study is that the data is derived from a single patient case study which makes generalized interpretation of biological implication of this data difficult. Furthermore, only a modest number of proteins was identified from these small tissue regions. As highly abundant proteins are identified preferentially in mass spectrometry analysis (70), the low number of proteins identified potentially masks differences in lower abundance proteins. The analysis of a larger cohort, coupled with the utilization of advanced sample preparation and mass spectrometry techniques to improve proteome coverage, holds the potential to build upon the data presented here and paint a clearer picture of the molecular landscape of precancerous lesions of serous cancers of the ovary and endometrium.

CONCLUSION

Here, we present the first proteomic investigation of precancerous lesions of the gynecological tract in a patient without advanced gynecological malignancy. Interpretation of the data is limited by the single case study and the modest number of proteins identified; however, we provide a foundation for further analysis of the molecular links between precancerous lesions and subsequent caner. In addition, we identified several metastasis-related proteins in precancerous tissues. The understanding that precancerous lesions of the female genital tract potentially possess metastatic potential raises many questions about when, where, and how these cancers develop. Though the early steps are not well understood, further proteomic analyses of gynecological precancerous lesions hold the potential to unravel the early temporal and molecular events underlying the development of these malignancies which, in turn, holds the potential to improve detection and treatment.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These can be found in the NCBI Gene Expression Omnibus. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (71) with the data set identifier PXD018538.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Committee of the Royal Adelaide Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Design of the work: MO, PH, GA, and MA. Acquisition, analysis, and interpretation of data: all authors. Drafting the work or revising it critically: all authors. All authors contributed to the article and approved the submitted version.

FUNDING

PH gratefully acknowledges the support of the National Collaborative Research Infrastructure Strategy (NCRIS) Bioplatforms Australia Node for Tissue Imaging Mass Spectrometry.

ACKNOWLEDGMENTS

We acknowledge the contribution of Lyron Winderbaum in analyzing the preliminary proteomics and genomics data gathered.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.523989/full#supplementary-material

REFERENCES

1. Amant F, Mirza MR, Koskas M, Creutzberg CL. Cancer of the corpus uteri. *Int J Gynaecol Obstet* (2015) 131(Suppl 2):S96–104. doi: 10.1016/j.ijgo.2015.06.005
2. Costas L, Frias-Gomez J, Guardiola M, Benavente Y, Pineda M, Pavon MA, et al. New perspectives on screening and early detection of endometrial cancer. *Int J Cancer* (2019) 145(12):3194–206. doi: 10.1002/ijc.32514
3. Urick ME, Bell DW. Clinical actionability of molecular targets in endometrial cancer. *Nat Rev Cancer* (2019) 19(9):510–21. doi: 10.1038/s41568-019-0177-x
4. Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, et al. Integrated genomic characterization of endometrial carcinoma. *Nature* (2013) 497(7447):67–73. doi: 10.1038/nature12113
5. Charo LM, Plaxe SC. Recent advances in endometrial cancer: a review of key clinical trials from 2015 to 2019. *F1000Res* (2019) 8. doi: 10.12688/f1000research.17408.1
6. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* (2015) 136(5):E359–86. doi: 10.1002/ijc.29210
7. Arnold M, Pandeya N, Byrnes G, Renehan PAG, Stevens GA, Ezzati PM, et al. Global burden of cancer attributable to high body-mass index in 2012: a population-based study. *Lancet Oncol* (2015) 16(1):36–46. doi: 10.1016/S1470-2045(14)71123-4
8. Soslow RA, Bissonnette JP, Wilton A, Ferguson SE, Alektiar KM, Duska LR, et al. Clinicopathologic analysis of 187 high-grade endometrial carcinomas of different histologic subtypes: similar outcomes belie distinctive biologic differences. *Am J Surg Pathol* (2007) 31(7):979–87. doi: 10.1097/PAS.0b013e31802ee494
9. Kelly MG, O'Malley D, Hui P, McAlpine J, Dziura J, Rutherford TJ, et al. Patients with uterine papillary serous cancers may benefit from adjuvant platinum-based chemoradiation. *Gynecol Oncol* (2004) 95(3):469–73. doi: 10.1016/j.ygyno.2004.08.030
10. Olawaiye AB, Rauh-Hain JA, Withiam-Leitch M, Rueda B, Goodman A, del Carmen MG. Utility of pre-operative serum CA-125 in the management of uterine papillary serous carcinoma. *Gynecol Oncol* (2008) 110(3):293–8. doi: 10.1016/j.ygyno.2008.05.027

11. Hamilton CA, Cheung MK, Osann K, Chen L, Teng NN, Longacre TA, et al. Uterine papillary serous and clear cell carcinomas predict for poorer survival compared to grade 3 endometrioid corpus cancers. *Br J Cancer* (2006) 94 (5):642–6. doi: 10.1038/sj.bjc.6603012
12. Zheng W, Liang SX, Yu H, Rutherford T, Chambers SK, Schwartz PE. Endometrial glandular dysplasia: a newly defined precursor lesion of uterine papillary serous carcinoma. Part I: morphologic features. *Int J Surg Pathol* (2004) 12(3):207–23. doi: 10.1177/106689690401200302
13. Liang SX, Chambers SK, Cheng L, Zhang S, Zhou Y, Zheng W. Endometrial glandular dysplasia: a putative precursor lesion of uterine papillary serous carcinoma. Part II: molecular features. *Int J Surg Pathol* (2004) 12(4):319–31. doi: 10.1177/106689690401200405
14. Zheng W, Xiang L, Fadare O, Kong B. A proposed model for endometrial serous carcinogenesis. *Am J Surg Pathol* (2011) 35(1):e1–14. doi: 10.1097/PAS.0b013e318202772e
15. Jia L, Liu Y, Yi X, Miron A, Crum CP, Kong B, et al. Endometrial glandular dysplasia with frequent p53 gene mutation: a genetic evidence supporting its precancer nature for endometrial serous carcinoma. *Clin Cancer Res* (2008) 14 (8):2263–9. doi: 10.1158/1078-0432.CCR-07-4837
16. Sherman ME, Bitterman P, Rosenshein NB, Delgado G, Kurman RJ. Uterine serous carcinoma. A morphologically diverse neoplasm with unifying clinicopathologic features. *Am J Surg Pathol* (1992) 16(6):600–10. doi: 10.1097/00000478-199206000-00008
17. Sherman ME, Bur ME, Kurman RJ. p53 in endometrial cancer and its putative precursors: evidence for diverse pathways of tumorigenesis. *Hum Pathol* (1995) 26(11):1268–74. doi: 10.1016/0046-8177(95)90204-X
18. Wheeler DT, Bell KA, Kurman RJ, Sherman ME. Minimal uterine serous carcinoma: diagnosis and clinicopathologic correlation. *Am J Surg Pathol* (2000) 24(6):797–806. doi: 10.1097/00000478-200006000-00004
19. Hui P, Kelly M, O'Malley DM, Tavassoli F, Schwartz PE. Minimal uterine serous carcinoma: a clinicopathological study of 40 cases. *Mod Pathol* (2005) 18(1):75–82. doi: 10.1038/modpathol.3800271
20. Tolcher MC, Swisher EM, Medeiros F, Lima JF, Hilderbrand JL, Donovan JL, et al. Characterization of precursor lesions in the endometrium and fallopian tube epithelium of early-stage uterine serous carcinoma. *Int J Gynecol Pathol* (2015) 34(1):57–64. doi: 10.1097/PGP.0000000000000109
21. Fader AN, Boruta D, Olawaiye AB, Gehrig PA. Uterine papillary serous carcinoma: epidemiology, pathogenesis and management. *Curr Opin Obstet Gynecol* (2010) 22(1):21–9. doi: 10.1097/GCO.0b013e328334d8a3
22. Sovak MA, Hensley ML, Dupont J, Ishill N, Alektiar KM, Abu-Rustum N, et al. Paclitaxel and carboplatin in the adjuvant treatment of patients with high-risk stage III and IV endometrial cancer: a retrospective study. *Gynecol Oncol* (2006) 103(2):451–7. doi: 10.1016/j.ygyno.2006.03.019
23. Bowtell DD. The genesis and evolution of high-grade serous ovarian cancer. *Nat Rev Cancer* (2010) 10(11):803–8. doi: 10.1038/nrc2946
24. George SH, Shaw P. BRCA and Early Events in the Development of Serous Ovarian Cancer. *Front Oncol* (2014) 4:5. doi: 10.3389/fonc.2014.00005
25. Piek JM, van Diest PJ, Zweemer RP, Jansen JW, Poort-Keesom RJ, Menko FH, et al. Dysplastic changes in prophylactically removed Fallopian tubes of women predisposed to developing ovarian cancer. *J Pathol* (2001) 195 (4):451–6. doi: 10.1002/path.1000
26. Crum CP, Drapkin R, Kindelberger D, Medeiros F, Miron A, Lee Y. Lessons from BRCA: the tubal fimbria emerges as an origin for pelvic serous cancer. *Clin Med Res* (2007) 5(1):35–44. doi: 10.3121/cmr.2007.702
27. Crum CP, Herfs M, Ning G, Bijron JG, Howitt BE, Jimenez CA, et al. Through the glass darkly: intraepithelial neoplasia, top-down differentiation, and the road to ovarian cancer. *J Pathol* (2013) 231 (4):402–12. doi: 10.1002/path.4263
28. Folkins AK, Jarboe EA, Saleemuddin A, Lee Y, Callahan MJ, Drapkin R, et al. A candidate precursor to pelvic serous cancer (p53 signature) and its prevalence in ovaries and fallopian tubes from women with BRCA mutations. *Gynecol Oncol* (2008) 109(2):168–73. doi: 10.1016/j.ygyno.2008.01.012
29. Kuhn E, Kurman RJ, Vang R, Sehdev AS, Han G, Soslow R, et al. TP53 mutations in serous tubal intraepithelial carcinoma and concurrent pelvic high-grade serous carcinoma–evidence supporting the clonal relationship of the two lesions. *J Pathol* (2012) 226(3):421–6. doi: 10.1002/path.3023
30. Lee Y, Miron A, Drapkin R, Nucci MR, Medeiros F, Saleemuddin A, et al. A candidate precursor to serous carcinoma that originates in the distal fallopian tube. *J Pathol* (2007) 211(1):26–35. doi: 10.1002/path.2091
31. Salvador S, Rempel A, Soslow RA, Gilks B, Huntsman D, Miller D. Chromosomal instability in fallopian tube precursor lesions of serous carcinoma and frequent monoclonality of synchronous ovarian and fallopian tube mucosal serous carcinoma. *Gynecol Oncol* (2008) 110(3):408–17. doi: 10.1016/j.ygyno.2008.05.010
32. Labidi-Galy SI, Papp E, Hallberg D, Niknafs N, Adleff V, Noe M, et al. High grade serous ovarian carcinomas originate in the fallopian tube. *Nat Commun* (2017) 8(1):1093. doi: 10.1038/s41467-017-00962-1
33. Xiang L, Rong G, Zhao J, Wang Z, Shi F. Identification of candidate genes associated with tubal origin of high-grade serous ovarian cancer. *Oncol Lett* (2018) 15(5):7769–75. doi: 10.3892/ol.2018.8346
34. Perets R, Wyant GA, Muto KW, Bijron JG, Poole BB, Chin KT, et al. Transformation of the fallopian tube secretory epithelium leads to high-grade serous ovarian cancer in Brca;Tp53;Pten models. *Cancer Cell* (2013) 24 (6):751–65. doi: 10.1016/j.ccr.2013.10.013
35. Zhai Y, Wu R, Kuick R, Sessine MS, Schulman S, Green M, et al. High-grade serous carcinomas arise in the mouse oviduct via defects linked to the human disease. *J Pathol* (2017) 243(1):16–25. doi: 10.1002/path.4927
36. Bijron JG, Seldenrijk CA, Zweemer RP, Lange JG, Verheijen RH, van Diest PJ. Fallopian tube intraluminal tumor spread from noninvasive precursor lesions: a novel metastatic route in early pelvic carcinogenesis. *Am J Surg Pathol* (2013) 37(8):1123–30. doi: 10.1097/PAS.0b013e318282da7f
37. Kim J, Coffey DM, Creighton CJ, Yu Z, Hawkins SM, Matzuk MM. High-grade serous ovarian cancer arises from fallopian tube in a mouse model. *Proc Natl Acad Sci USA* (2012) 109(10):3921–6. doi: 10.1073/pnas.1117135109
38. Lim D, Oliva E. Precursors and pathogenesis of ovarian carcinoma. *Pathology* (2013) 45(3):229–42. doi: 10.1097/PAT.0b013e32835f2264
39. Marquez RT, Baggerly KA, Patterson AP, Liu J, Broaddus R, Frumovitz M, et al. Patterns of gene expression in different histotypes of epithelial ovarian cancer correlate with those in normal fallopian tube, endometrium, and colon. *Clin Cancer Res* (2005) 11(17):6116–26. doi: 10.1158/1078-0432.CCR-04-2509
40. Roelofsen T, van Kempen LC, van der Laak JA, van Ham MA, Bulten J, Massuger LF. Concurrent endometrial intraepithelial carcinoma (EIC) and serous ovarian cancer: can EIC be seen as the precursor lesion? *Int J Gynecol Cancer* (2012) 22(3):457–64. doi: 10.1097/IGC.0b013e3182434a81
41. Mittal P, Klingler-Hoffmann M, Arentz G, Winderbaum L, Kaur G, Anderson L, et al. Annexin A2 and alpha actinin 4 expression correlates with metastatic potential of primary endometrial cancer. *Biochim Biophys Acta Proteins Proteom* (2017) 1865(7):846–57. doi: 10.1016/j.bbapap.2016.10.010
42. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods* (2009) 6(5):359–62. doi: 10.1038/nmeth.1322
43. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* (2011) 10(4):1794–805. doi: 10.1021/pr101065j
44. Yoshihara K, Tajima A, Komata D, Yamamoto T, Kodama S, Fujiwara H, et al. Gene expression profiling of advanced-stage serous ovarian cancers distinguishes novel subclasses and implicates ZEB2 in tumor progression and prognosis. *Cancer Sci* (2009) 100(8):1421–8. doi: 10.1111/j.1349-7006.2009.01204.x
45. Hendrix ND, Wu R, Kuick R, Schwartz DR, Fearon ER, Cho KR. Fibroblast growth factor 9 has oncogenic activity and is a downstream target of Wnt signaling in ovarian endometrioid adenocarcinomas. *Cancer Res* (2006) 66 (3):1354–62. doi: 10.1158/0008-5472.CAN-05-3694
46. Day RS, McDade KK, Chandran UR, Lisovich A, Conrads TP, Hood BL, et al. Identifier mapping performance for integrating transcriptomics and proteomics experimental results. *BMC Bioinf* (2011) 12:213. doi: 10.1186/1471-2105-12-213
47. Li Z, Huang C, Bai S, Pan X, Zhou R, Wei Y, et al. Prognostic evaluation of epidermal fatty acid-binding protein and calcyphosine, two proteins implicated in endometrial cancer using a proteomic approach. *Int J Cancer* (2008) 123(10):2377–83. doi: 10.1002/ijc.23808
48. Wen KC, Sung PL, Chou YT, Pan CM, Wang PH, Lee OK, et al. The role of EpCAM in tumor progression and the clinical prognosis of endometrial

carcinoma. *Gynecol Oncol* (2018) 148(2):383–92. doi: 10.1016/j.ygyno.2017.11.033
49. Tayama S, Motohara T, Narantuya D, Li C, Fujimoto K, Sakaguchi I, et al. The impact of EpCAM expression on response to chemotherapy and clinical outcomes in patients with epithelial ovarian cancer. *Oncotarget* (2017) 8 (27):44312–25. doi: 10.18632/oncotarget.17871
50. Erickson BK, Conner MG, Landen CN Jr. The role of the fallopian tube in the origin of ovarian cancer. *Am J Obstet Gynecol* (2013) 209(5):409–14. doi: 10.1016/j.ajog.2013.04.019
51. Blaustein A. Peritoneal mesothelium and ovarian surface cells–shared characteristics. *Int J Gynecol Pathol* (1984) 3(4):361–75.
52. Reece-Smith AM, Parsons SL, Watson SA. Oesophago-Gastric Cancer. In: Bologna M, editor. *Biotargets of Cancer in Current Clinical Practice. Current Clinical Pathology*. Humana Press (2012). p. 221–44. doi: 10.1007/978-1-61779-615-9_8
53. Szotek PP, Chang HL, Brennand K, Fujino A, Pieretti-Vanmarcke R, Lo Celso C, et al. Normal ovarian surface epithelial label-retaining cells exhibit stem/progenitor cell characteristics. *Proc Natl Acad Sci USA* (2008) 105(34):12469–73. doi: 10.1073/pnas.0805012105
54. Nobumoto A, Nagahara K, Oomizu S, Katoh S, Nishi N, Takeshita K, et al. Galectin-9 suppresses tumor metastasis by blocking adhesion to endothelium and extracellular matrices. *Glycobiology* (2008) 18(9):735–44. doi: 10.1093/glycob/cwn062
55. Fujihara S, Mori H, Kobara H, Rafiq K, Niki T, Hirashima M, et al. Galectin-9 in cancer therapy. *Recent Pat Endocr Metab Immune Drug Discovery* (2013) 7 (2):130–7. doi: 10.2174/1872148113070200006
56. Cui X, Song B, Hou L, Wei Z, Tang J. High expression of osteoglycin decreases the metastatic capability of mouse hepatocarcinoma Hca-F cells to lymph nodes. *Acta Biochim Biophys Sin (Shanghai)* (2008) 40(4):349–55. doi: 10.1111/j.1745-7270.2008.00392.x
57. Wang L, Zhang F, Cui JY, Chen L, Chen YT, Liu BW. CAFs enhance paclitaxel resistance by inducing EMT through the IL6/JAK2/STAT3 pathway. *Oncol Rep* (2018) 39(5):2081–90. doi: 10.3892/or.2018.6311
58. Yoshida R, Kimura N, Harada Y, Ohuchi N. The loss of E-cadherin, alpha- and beta-catenin expression is associated with metastasis and poor prognosis in invasive breast cancer. *Int J Oncol* (2001) 18(3):513–20. doi: 10.3892/ijo.18.3.513
59. Beavon IR. The E-cadherin-catenin complex in tumour metastasis: structure, function and regulation. *Eur J Cancer* (2000) 36(13 Spec No):1607–20. doi: 10.1016/S0959-8049(00)00158-1
60. Kobel M, Kalloger S, Boyd NJPM. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med* (2008) 5: (12):e232. doi: 10.1371/journal.pmed.0050232
61. Kuechlin S, Schoels M, Slanchev K, Lassmann S, Walz G, Yakulov TA. EpCAM controls morphogenetic programs during zebrafish pronephros development. *Biochem Biophys Res Commun* (2017) 487(2):209–15. doi: 10.1016/j.bbrc.2017.04.035
62. Lu TY, Lu RM, Liao MY, Yu J, Chung CH, Kao CF, et al. Epithelial cell adhesion molecule regulation is associated with the maintenance of the undifferentiated phenotype of human embryonic stem cells. *J Biol Chem* (2010) 285(12):8719–32. doi: 10.1074/jbc.M109.077081
63. Clement S, Dumont JE, Schurmans S. Loss of calcyphosine gene expression in mouse and other rodents. *Biochem Biophys Res Commun* (1997) 232(2):407–13. doi: 10.1006/bbrc.1997.6297
64. Kusumoto T, Kodama J, Seki N, Nakamura K, Hongo A, Hiramatsu Y. Clinical significance of syndecan-1 and versican expression in human epithelial ovarian cancer. *Oncol Rep* (2010) 23(4):917–25. doi: 10.3892/or_00000715
65. Yamaguchi K, Mandai M, Oura T, Matsumura N, Hamanishi J, Baba T, et al. Identification of an ovarian clear cell carcinoma gene signature that reflects inherent disease biology and the carcinogenic processes. *Oncogene* (2010) 29 (12):1741–52. doi: 10.1038/onc.2009.470
66. Yamamoto S, Tsuda H, Takano M, Iwaya K, Tamai S, Matsubara O. PIK3CA mutation is an early event in the development of endometriosis-associated ovarian clear cell adenocarcinoma. *J Pathol* (2011) 225(2):189–94. doi: 10.1002/path.2940
67. Levanon K, Ng V, Piao HY, Zhang Y, Chang MC, Roh MH, et al. Primary ex vivo cultures of human fallopian tube epithelium as a model for serous ovarian carcinogenesis. *Oncogene* (2010) 29(8):1103–13. doi: 10.1038/onc.2009.402
68. Eckert MA, Coscia F, Chryplewicz A, Chang JW, Hernandez KM, Pan S, et al. Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. *Nature* (2019) 569(7758):723–8. doi: 10.1038/s41586-019-1173-8
69. Koscielny S, Tubiana M, Le MG, Valleron AJ, Mouriesse H, Contesso G, et al. Breast cancer: relationship between the size of the primary tumour and the probability of metastatic dissemination. *Br J Cancer* (1984) 49(6):709–15. doi: 10.1038/bjc.1984.112
70. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* (2003) 422(6928):198–207. doi: 10.1038/nature01511
71. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* (2019) 47 (D1):D442–50. doi: 10.1093/nar/gky1106

Conflict of Interest: Author FW was employed by the company Clinpath.

The remaining 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 © 2020 Acland, Arentz, Mussared, Whitehead, Hoffmann, Klingler-Hoffmann and Oehler. 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 Oncology

ORIGINAL RESEARCH
published: 22 December 2020
doi: 10.3389/fonc.2020.543947

A Protein Microarray-Based Investigation of Cerebrospinal Fluid Reveals Distinct Autoantibody Signature in Low and High-Grade Gliomas

*Nikita Gahoi[1,2], Parvez Syed[1,3], Saket Choudhary[4,5], Sridhar Epari[6], Aliasgar Moiyadi[7], Santosh G. Varma[8,9], Mayuri N. Gandhi[2] and Sanjeeva Srivastava[1]**

[1] Wadhwani Research Center for Biosciences and Bioengineering, Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India, [2] Centre for Research in Nanotechnology and Sciences, Indian Institute of Technology Bombay, Mumbai, India, [3] Inme Oy, Turku, Finland, [4] Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India, [5] Molecular and Computational Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA, United States, [6] Department of Pathology, Tata Memorial Centre, Mumbai, India, [7] Neurosurgical Services, Department of Surgical Oncology, Tata Memorial Centre and Homi Bhabha National Institute, Mumbai, India, [8] Deptartment of Biochemistry, Grant Govt. Medical College and Sir JJ Group of Hospitals, Mumbai, India, [9] Department of Biochemistry, BJ Medical College and Sassoon Hospital, Pune, India

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
David D. Gonda,
University of California, San Diego, United States
Jonathan M. Blackburn,
University of Cape Town, South Africa

***Correspondence:**
Sanjeeva Srivastava
sanjeeva@iitb.ac.in

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 27 March 2020
Accepted: 23 October 2020
Published: 22 December 2020

Citation:
Gahoi N, Syed P, Choudhary S, Epari S, Moiyadi A, Varma SG, Gandhi MN and Srivastava S (2020) A Protein Microarray-Based Investigation of Cerebrospinal Fluid Reveals Distinct Autoantibody Signature in Low and High-Grade Gliomas. Front. Oncol. 10:543947. doi: 10.3389/fonc.2020.543947

Gliomas are one of the most aggressive primary brain tumors arising from neural progenitor cells. Delayed diagnosis, invasive biopsy, and diagnostic challenges stems the need for specific, minimally-invasive, and early diagnostic biomarkers. Tumor-associated (TA) autoantibodies are measurable in the biofluids long before the onset of the symptoms, suggesting their role in early diagnosis and clinical management of the patients. In the current study, cerebrospinal fluid (CSF) samples from patients with low-grade glioma (LGG) and the *Glioblastoma multiforme* (GBM) that characterizes advanced disease were compared with healthy control samples to identify putative TA autoantibodies, using protein microarrays. The CSF samples from LGGs (n = 10), GBM (n = 7) were compared with the control CSF samples (n = 6). Proteins showing significant antigenic response were cross-verified. Proteins NOL4 (a cancer-testis antigen) and KALRN showed an antigenic response in the CSF of GBM patients, whereas, UTP4 and CCDC28A showed an antigenic response in low grade gliomas when compared with the control samples. TA autoantibodies identified in this study from the CSF of the patients could supplement current screening modalities. Further validation of these TA autoantibodies on a larger clinical cohort could provide cues towards relevance of these proteins in early diagnosis of the disease.

Keywords: cancer biomarkers, glioma, cerebrospinal fluid, protein microarray, *Glioblastoma multiforme*, autoantibodies

INTRODUCTION

Gliomas are one of the most aggressive primary brain tumors characterized by high morbidity and mortality rates due to their localization, invasive, and heterogeneous nature (1). Based on the histological characteristics, gliomas were traditionally classified into astrocytic, oligodendroglial, or ependymal tumor and were further assigned WHO grades I–IV, indicating the degree of invasiveness and malignancy. The advancement in genomic, transcriptomic and epigenetic profiling techniques has improvised the molecular understanding of gliomas and led to reclassification of these tumors based upon tumor histories, response to treatment and clinical outcomes (2). The 2016 WHO classification, grouped the diffused low-grade gliomas (WHO grade II) and intermediate-grade gliomas (WHO grade III) as low-grade gliomas (LGGs). These LGGs are further sub-divided based on the molecular markers like mutations in IDH, ATRX, TP53, and co-deletion of 1p and 19q arms of chromosomes (3). Pilocytic astrocytomas (WHO Grade I), the most common type of glioma in children, are molecularly distinct from adult gliomas. These are characterized by favorable prognosis, circumscribed growth and frequently carry BRAF gene mutations or fusion. Grade IV gliomas also known as *Glioblastoma multiforme* (GBM) are most aggressive among gliomas and characterized by cells having high mitotic rates, nuclear atypia with adjoining areas of new vessel formation and necrosis. GBMs are highly infiltrative, favor growth surrounding glial cells by suppressing immune response. GBMs are further classified into primary and secondary GBMs. Primary GBMs develop *de novo* and are highly aggressive with a survival period of less than 2 years, whereas secondary GBMs evolves gradually from lower grade gliomas and have a better prognosis than primary GBMs. Primary and secondary GBMs are histopathologically indistinguishable, however they harbor different molecular alterations (4).

Tumorigenesis is a result of accumulation of mutations in cells over time leading to production of aberrant biomolecules that may be antigenic. These atypical biomolecules are commonly referred to as tumor associated (TA) antigens and evoke immunological response resulting in production of antibodies against self-proteins, also known as TA autoantibodies. TA autoantibodies have characteristics of antibodies like specificity, production in large quantities, long half-lives and circulation in biofluids (5) and have been reported to be diagnosed five years before the development of clinical symptoms (6, 7) making them promising biomarker candidates in cancers. Till date, several TA autoantibodies have been reported in sera of cancer patients. However, the presence of blood-brain barrier in central nervous system (CNS) restricts the entry of these autoantibodies into the patient sera. The cerebrospinal fluid in the CNS acts as an amalgamation of the transported biological substances, the waste and toxic substances excreted from the brain. This makes the CSF an invaluable source of biomarkers for diagnosis, prognosis of the course of CNS disease, and also as a predictive biofluid in the presymptomatic state.

High density protein microarrays provide an important platform for antigen display (8) and allow assessment of thousands of tumor antigens simultaneously with a minimal requirement of biological samples. These arrays have been applied to study immune response against several auto-immune diseases and cancers. To screen the presence of TA autoantibody signature in CSF of the glioma patients, protein microarray (ProtoArray© V5.0) containing ~9,000 human recombinant proteins has been used. To the best of our knowledge this is the first pilot experiment where CSF has been used to screen the presence of TA autoantibodies in the glioma patients.

Abbreviations: BBB, Blood Brain Barrier; CCDC28A, Coiled-Coil domain containing 28A; CNS, Central Nervous System; CSF, Cerebrospinal Fluid; FDR, False Discovery Rate; GBM, *Glioblastoma Multiforme*; IDH, Isocitrate Dehydrogenase; KALRN, Kalirin; LGG, Low-Grade Glioma; NOL4, Nucleolar protein 4; TA, Tumor-associated Autoantibodies; UTP4, U3 small nucleolar RNA-associated protein 4 homolog.

MATERIALS AND METHODS

Patient CSF Sample Information

CSF samples used in this study were collected from Advanced Centre for Treatment Research and Education in Cancer (ACTREC), and Tata Memorial Hospital, Mumbai, India. The study was approved by the TMC-ACTREC-Institutional Review Board (ACTREC-TMC IEC No. 15). The samples were collected after prior written consent. CSF was collected from the radiologically verified glioma patients undergoing surgery. CSF samples from 17 individuals with gliomas and 6 control patients (without any history of any intracranial infection or surgery) were used for autoantibody screening. The aliquots of the collected samples were stored at -80°C until further use. The CSF samples used in the study were classified as low-grade glioma LGGs (n = 10), GBM (n = 7), and control subjects (n = 6). Detailed information about the samples is provided in **Supplementary Tables 1** and **2**.

Microarray Analysis

Human protein microarray (ProtoArray© V5.0), were purchased from Invitrogen, Carlsbad, CA. Each ProtoArray protein array contains around 9,000 GST-tagged full-length human proteins in duplicates. These full-length N-terminal GST fusion proteins are expressed using Baculovirus expression system and were purified under non-denaturing conditions to maintain the protein integrity and function. The purified proteins are then printed on an ultrathin layer of nitrocellulose coated glass slides under temperature and humidity-controlled environment. Each block on a ProtoArray slide contains positive (Alexa Flour Ab, Human IgG, Anti-human IgG, and V5 control protein) and negative (Buffer, BSA, and GST) control spotted in duplicates (**Figure 1B**).

Autoantibody Screening

The ProtoArray slides were blocked using SuperBlock (Pierce) at 4°C for 1 h with gentle shaking followed by washing with PBST buffer (4 x 5 min) at room temperature. Each slide was probed with single CSF sample and a dilution of 1:4 CSF in blocking buffer (25 µl of CSF in 75 µl of blocking buffer) was incubated onto the slides overnight using coverslips. The slides were then washed with

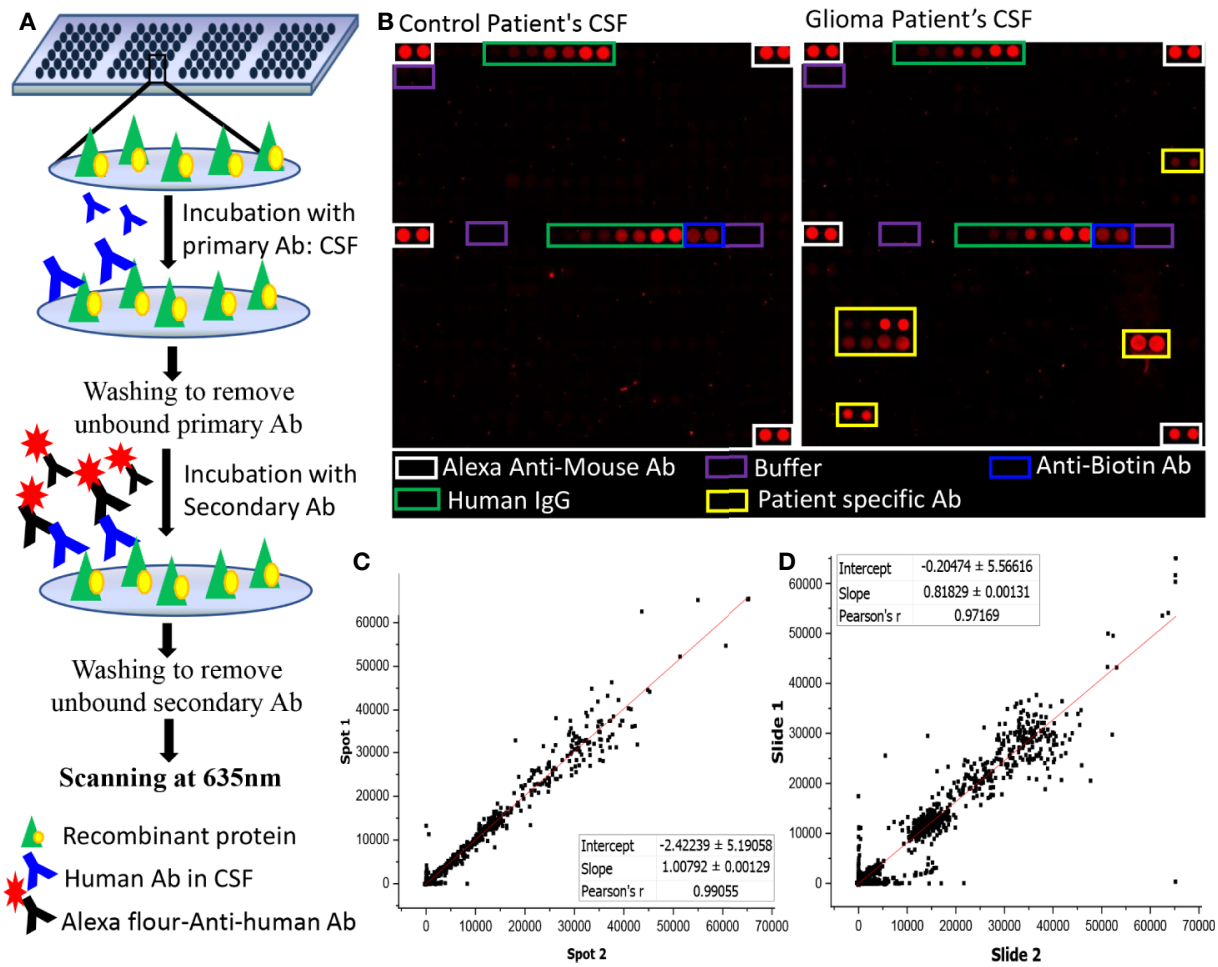


FIGURE 1 | Autoantibody screening using protoarray. **(A)** Schematic representation of the experimental workflow; **(B)** Image of two sub-arrays printed on a protoarray slide, left panel represents the location of control spots on the subarray and right panel represents the location of control spots on the subarray along with signal response due to patient specific Ab; **(C, D)** Graph representing **(C)** intra- chip reproducibility, spot intensities of duplicate protein spots from same array; **(D)** Inter-chip reproducibility of slides, intensities of protein spots from two different protoarrays processed using two different samples were plotted.

PBST buffer (4X5 mins), rinsed in distilled water and dried at 900 rpm for 2 min. A dilution of 1:5,000 of Cy5-labelled Goat Anti-Human Ab in blocking buffer was used as secondary Ab and the slides were incubated for an hour at room temperature on the shaker. After which the slides were washed with PBST (4 times at 5 min interval each), rinsed with distilled water and centrifuged at 4°C for 2 min at 900 rpm. Scanning of the dried slides was performed at 635 nm at 600 PMT gain using GenePix 4000B Microarray Scanner (Molecular Devices) (**Figure 1A**).

Statistical Analysis for the Identification of Autoantibody Signatures

Images obtained after scanning the microarray slides were processed using GenePix Pro7 (Molecular Devices) software. The acquired data, that is, the median pixel intensity values obtained for each protein spot was analyzed using two different methods. In the first method, median pixel intensity (F635) subtracted by local background (B635) for each spot was considered for analysis (median F635-B635). A base cut-off of 60 was set to replace all the negative values and values below 60. Sample specific normalization was performed by subtracting the median value of multiple buffer spot present in the array to the intensity of protein spots that are printed throughout the chip (9). The normalized signal intensities of all the duplicate spots were averaged and log2 transformed, this value was further used for statistical analysis using Metaboanalyst 4.0 (10). Two-tailed *t*-statistics was applied to identify the proteins with antigenic potential. Further, the p-values were adjusted for false discovery using Bonferroni correction. To cross-examine the data obtained, dot plots for all the significant proteins (raw p-value <0.005 and absolute FC >1.2) were plotted for different comparison using GraphPad Prism (Prism v6.0, GraphPad Software Inc., La Jolla CA).

In the second method, the data was pre-processed to adjust the technical errors between the arrays, thereby adjusting the differences that did not arise biologically. "Limma" package (11) made available as a Bioconductor package (12) for R programming

was used to pre-process the data. Pre-processing of the data was performed using the "neqc" function that performs both background correction and quantile normalization using a set of negative control genes. Pre-processed data was then analyzed to look for the proteins with antigenic potential. "Limma" package uses moderated *t*-statistic to test the null hypothesis that proteins are not differentially expressed between two conditions. To adjust for multiple hypothesis testing, we used, "Benjamini-Hochberg" (BH) correction. Statistically significant proteins with adjusted p-value <0.05 were sorted and the proteins with fold-change ≥ ± 1.5 were considered for further analysis. Autoantibody response for these proteins was manually verified for all the samples.

Gene Set Enrichment Analysis

To understand the functional, molecular and sub-cellular characteristics, gene set enrichment analysis of proteins showing antigenic characteristic (p-value <0.05 and abs FC >1.2) was performed using software like DAVID (DAVID Bioinformatics Resources 6.7) (13) and Protein ANalysis THrough Evolutionary Relationships (PANTHER) (14) system version 7 (http://www.pantherdb.org).

RESULTS

Quality Check of the Processed Slides

The intra-chip variation was evaluated by calculating the sample coefficient of variation (CV) between the duplicate spots using "CV" function of R programming. The CV were calculated for each slide using the raw pixel intensity (median F635-B635). The value of CV ranged from 0.14 to 0.31 with an average of 0.215 ± 0.04 (**Supplementary Table 3**). Further, a graphical representation of intra-chip and inter-chip correlation is given in **Figures 1C**, **D**. The intensities of the protein spots were checked for all slides showing lower intensities in the positive control spots. The protein spots showed comparable signal intensities with protein spot on other slides, hence, none of the samples were excluded from the study.

Autoantibody Signatures in Glioma Samples

The normalized data was subjected to moderated t-test and the obtained p-values were corrected using Bonferroni correction. Due to the stringent nature of the Boneferroni method, none of the proteins could pass the filtering criteria of adj. p-value <0.05. However, with a filtering criterion of raw p-value <0.05 and abs FC >1.2, a total of 15 proteins were found to show variable expression in the comparison of LGG vs. Control, of which 13 protein showed a positive fold changes while 2 proteins showed a negative fold change with respect to control samples (**Figure 2A**). The comparison of GBM with control samples yielded total 405 proteins, all of which showed a positive response in GBM samples (**Figure 2A**). Further, a cut-off of p-value less than 0.005 and abs FC >1.2, filtered a total of 33 and 3 proteins from GBM and LGG respectively when compared to control samples, list of these proteins with their fold changes and respective p-values is provided in **Supplementary Table 4**. Further, to investigate the segregating potential of these filtered proteins, partial least squares discriminant analysis (PLS-DA) and unsupervised clustering was performed. PLS-DA showed good segregation of the GBM samples from control and LGG samples whereas LGG and Control samples showed some overlap (**Figure 2B**). The unsupervised clustering revealed that the GBM samples clustered together but there was a significant overlap between LGG and Control samples (**Supplementary Figure 1**). These results revealed the presence of TA autoantibodies in the CSF of GBM samples is much higher as compared to low grade glioma samples. The data was further cross-validated by plotting the raw intensity values of the duplicate spots for all the proteins passing the filtering criteria of raw p-value <0.005 and FC >1.2 (**Supplementary Figure 2**).

In the other method, pre-processing of the data was performed using "neqc" function that performs both background correction and quantile normalization using the negative control spots. The pre-processed data seemed to be median centric, thereby excluding any possible technical variances and was used for further analysis. The list of differentially expressed proteins was further subjected to moderated t-statistic and "Benjamini-Hochberg" (BH) correction. The criteria of adj. p-value <0.05 and absolute FC >1.5 was applied and proteins passing the criteria were further considered to be showing an immunogenic response. In a comparison of low-grade glioma and control, 2 proteins viz., U3 small nucleolar RNA-associated protein 4 homolog (UTP4) and Coiled-Coil domain containing 28A (CCDC28A) showed a significant response. The comparison between GBM vs. control resulted in 2 proteins, namely Nucleolar protein 4/Cancer/Testis Antigen 125 (NOL4) and Kalirin (KALRN), showed an immunogenic response (**Table 1**). These proteins were also present in the list obtained from the first analysis and were further cross-validated manually by looking at the individual spots on each slide. Intensities of these significant proteins are represented in form of whisker's plot along with their respective protein spots on the protoarray slide in **Figure 2C**.

Gene Set Enrichment Analysis

Bioinformatics tools like Panther and DAVID were used to characterize the list of proteins obtained from the statistical analysis using Metaboanalyst. The sub-cellular localization of 85.65% of these proteins was cytoplasmic, organelle-based or were a part of macromolecular complex system, while only 14.00% of the protein had membrane-linked or extra-cellular matrix-based origin. Most of these proteins categorized into protein modifying enzyme (19.20%), nucleic acid binding protein (13.60%), and metabolite interconversion enzyme (10.20%). Some of these proteins were associated with angiogenesis, inflammation, B cell activation pathway, signaling pathway, FGF signaling pathway, PDGF signaling pathways, Wnt and Notch signaling pathway, Ras pathway, p53 pathway and EGFR signaling pathway.

DISCUSSION

Gliomas are neoplasms arising from neuroepithelial tissues. Molecular alteration in the neoplasm results in production of anomalous biomolecules that evokes immune response leading

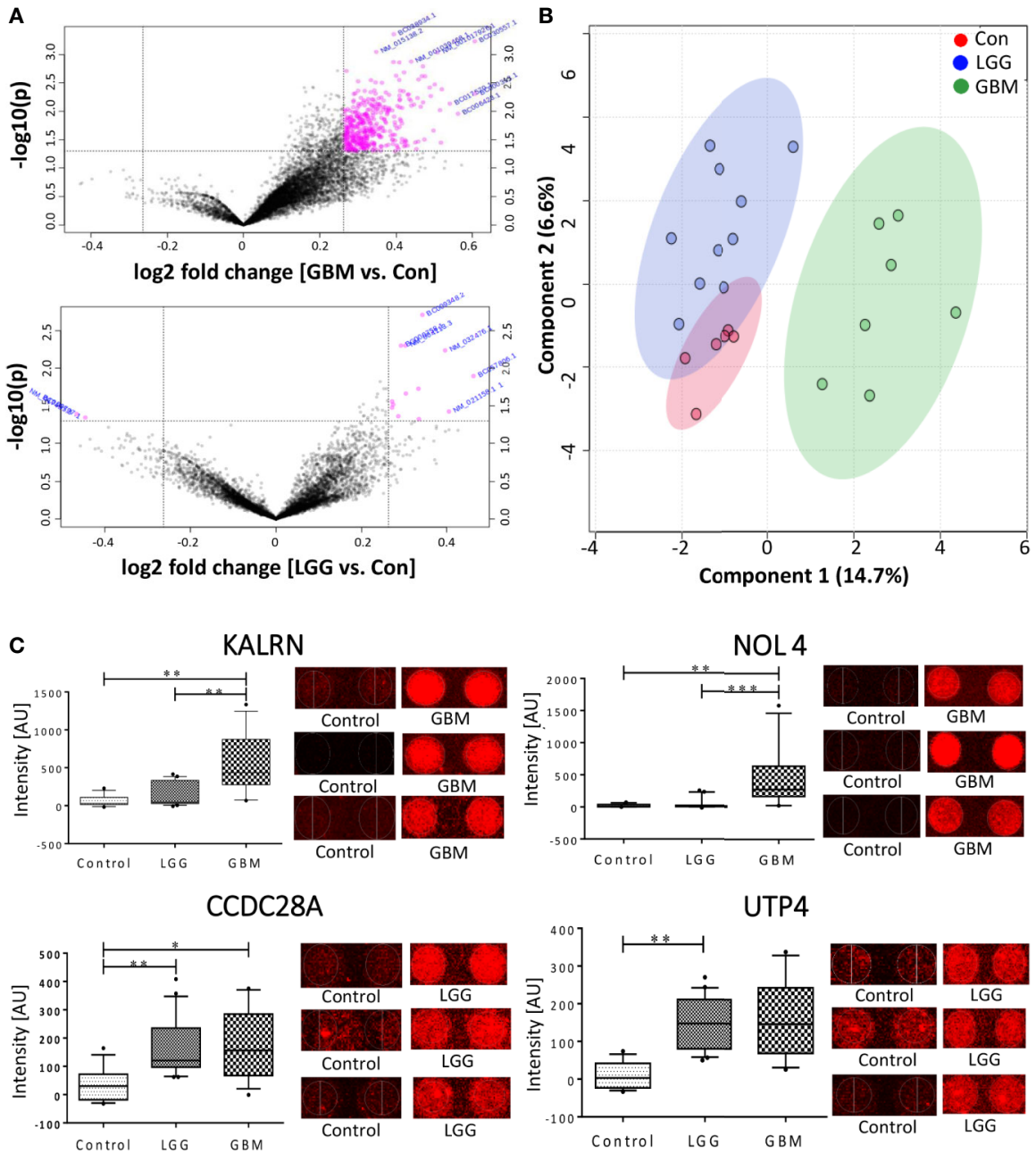


FIGURE 2 | Autoantibody response in CSF of Glioma patients. **(A)** Volcano plot signifying the potentially antigenic protein in GBM and LGG samples **(B)** SPLSDA plot representing the segregation of control, LGG and GBMs, GBM showed a good segregation from control and low grade glioma; **(C)** Whiskers plot of some of the significant proteins in GBM and LGG and their respective spots on the protoarray slide. (*indicates $0.01 < p < 0.05$, **indicates $0.001 < p < 0.01$, ***indicates $0.0001 < p < 0.001$).

to production of TA autoantibodies, against self-proteins. TA autoantibodies are highly specific and are produced against modified or amplified tumor marker, with a half-life of up to 30 days in circulation and can be quantified using routinely used platforms in the clinics. These TA autoantibodies majorly belong to IgG class of immunoglobulin and have been reported to be present in the serum of cancer patients. A study reported that the rate of diagnosis of organ-specific autoimmune neurologic disorders was clinically pertinent when paraneoplastic autoantibodies were detected in both serum and CSF of the patients with clinical suspicion (15). The presence of BBB makes CSF an invaluable source of biomarkers for diagnosis, prognosis of the course of CNS disease, and also as a predictive biofluid in the presymptomatic state.

TABLE 1 | List of significant proteins against which TA autoantibodies were observed in low grade glioma and *glioblastoma multiforme* samples.

Gene ID	Protein name	Uniprot ID	Log FC	FC	p-value	Adj. p-value
Low Grade Glioma vs. Control						
BC009348.2	U3 small nucleolar RNA-associated protein 4 homolog	UTP4	0.60	1.516	3.36E-05	0.033
BC000758.1	Coiled-Coil domain containing 28A	CCDC28A	0.619	1.536	6.80E-05	0.046
Glioblastoma Multiforme vs. Control						
BC000313.1	Nucleolar Protein 4	NOL4	1.356	2.56	5.78E-07	0.003
NM_007064.1	Kalirin	KALRN	1.264	2.40	6.19E-07	0.003

In the current study, we have used protein microarrays, containing ~9,000 recombinant proteins to detect TA autoantibody signatures in the CSF of glioma patients. The samples size used in the study was low, which increases the risk of error, therefore, the only proteins that could pass the adj p-value cut-off of less than 0.05 and fold change greater than 1.5 were considered and have been discussed in details. From the statistical analysis, two proteins viz., Nucleolar Protein 4 (NOL4) and Kalirin (KALRN) showed a significant antigenic response in GBM patients when compared to control samples. *NOL4* is a nucleolar protein encoding gene with a predominant expression in brain and testis. Aberrant methylation of CpG islands in the *NOL4* gene promoter has been reported to be associated with 85% of the cervical cancer patients (16) and 91% of head and neck squamous cell carcinoma (HNSCC) (17). An increased expression of *NOL4* has been reported in prostate cancer patients (18) and was reported to be significantly associated with the aggressiveness of the disease (19). Stangeland et al., reported *NOL4* to be one of the 20 aberrantly expressed gene in GBMs (20). KALRN is a member of Dbl family with two unique Rho guanine nucleotide exchange factors (GEFs) and multiple spectrin-like domains. mRNA expression profiling of *KALRN* gene revealed different isoforms of Kalirin produced through alternate splicing; its expression is restricted to CNS with the highest expression seen in cerebral cortex and hippocampus. Kalirin-7 is the most abundant isoform present in adult brain that regulates maturation and maintenance of dendritic spine, post-synaptic actin dynamics, axon extension and activity dependent plasticity. Mains et al., reported over-expression of Kalirin-8 (another isoform of Kalirin) in Chinese hamster ovary cells and AtT-20 cells that resulted in rearrangement of actin cytoskeleton (21). Greenman et al., performed a systemic genome sequence profiling of 518 kinases to identify possible mutations across 210 diverse human cancers. In this study, TRIO, a paralog of Kalirin, was also sequenced. Nine mutations were found in TRIO gene of which 6 mutations were present in the catalytic domains of TRIO. An insertion mutation of A156InsP (C469-470CCC) was found in glioma samples (22).

UTP4 and CCDC28A showed significant autoantigenic response in LGGs when compared with control sample. hUTP4/CIRH1A is a human ribosome biogenesis factor which is a part of t-UTP subcomplex of ribosomal small subunit (SSU) processome required for the maturation of 18S rRNA (23, 24). Homologous missense mutation (R565W) at the C-terminus of the CIRH1A has been reported in all the patients with North American Indian Childhood Cirrhosis (NAIC) (25). *In silico* analysis reported over-expression of CIRH1A in colon and para-rectum adenocarcinomas. Knock-down of CIRH1A in RKO CRC cell lines resulted in increased apoptosis, suppressed cell proliferation and cell arrest at G1 phase (26). Yu et al., reported that CIRH1A, interacts with Cirip/HIVEP and leads to up-regulation of NF-κB element (27), an active player in human cancers (28). *CCDC28A* encodes for coiled-coil domain containing protein and is a known translocation partner of nucleoporin 98 (NUP98) in T cell acute lymphoblastic leukaemia (T-ALL). Petit et al., transduced *NUP98-CCDC28A* in primary murine hematopoietic progenitor cells and reported that *NUP98-CCDC28A* promotes self-renewable potential and proliferative capacity of myeloid progenitors through an alternative leukemogenic mechanism (29).

The gene set enrichment analysis showed that only 14.00% of proteins were associated with cell junction, membrane, extracellular matrix or extracellular region. The remaining 85.65% of the proteins were associated with cytoplasm, organelle or were a part of macromolecular complex (**Supplementary Table 5**). Contrary to the serum antibodies that are produced predominantly against extra-cellular protein, TA autoantibodies are produced against intra-cellular markers. Reuschenbach et al., summarized in their review that the sub-cellular distribution of the proteins against which the TA autoantibodies is produced is 42% cytoplasmic, 26.1% nucleus associated, 21.4% membrane bound, and only 10.3% extracellular (30). Several studies have reported that the TA autoantibodies epitope could be highly conserved and bind to the functional site of the proteins inhibiting the function of antigen (31–33). Therefore, TA autoantibodies could be immunological biomarkers produced against aberrant cellular mechanism related to tumorigenesis (5, 34).

The results from our findings suggest the presence of TA autoantibodies in the CSF of glioma patients. LGGs and GBM showed autoantibody response against different set of proteins, which can be well corroborated with their distinct molecular characteristics. An enhanced presence of TA autoantibodies was observed in GBM samples as compared to the low-grade gliomas. These signature molecules can be quantified in biofluids using routine clinical platforms like ELISA and could aid in early diagnosis of disease. This was an exploratory study designed to detect the presence of TA autoantibodies in CSF of the cancer

patients. Nevertheless, the small sample size used in the study possess the risk of increased rate of error, therefore the findings from the study needs to be further validated on the larger cohort.

CONCLUSION

Gliomas are one of the most aggressive brain tumors and they are diagnosed at an advanced stage. Therefore, an early identification of disease with less invasive methods would help in improvising the morbidity and mortality rates. Mutations in neoplasm lead to production of atypical biomolecules that evokes immune response and leads to production of TA autoantibodies. TA autoantibodies are detectable in biofluids and their quantification platforms are already in clinical use. We investigated the presence of TA autoantibodies in the CSF of glioma patients to detect the pre-symptomatic markers indicative of diseased state. Most of the CSF samples procured from GBM patients showed the presence of TA autoantibodies against NOL4 and KALRN while low grade glioma samples showed an antigenic response against UTP4 and CCDC28A. Considering the fact that the sample size used in the study was small, these finding needs to be validated on a larger population. Detecting TA autoantibodies in cancer patients could aid in early diagnosis of disease and might provide useful insight into neoplasm initiation.

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 studies involving human participants were reviewed and approved by TMH-IRB. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NG and SS conceived the idea. NG and PS performed microarray experiments. SE, AM, and SV provided clinical samples and clinical details. NG, SC, and MG performed analysis. NG and SS contributed in manuscript preparation. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by Department of Biotechnology (DBT) (No. BT/PR14359/MED/30/916/2010) and (BT/PR4599/BRB/10/1042/2012), government of India, MHRD-UAY Project, Uchhatar Avishkar Yojana, project #34_IITB (2016) to SS and intramural funding from Tata Memorial Centre (DAE-CTC grant to AM). NG was supported by Council of Scientific and Industrial Research (CSIR) Senior Research Fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.543947/full#supplementary-material

REFERENCES

1. Rajesh Y, Pal I, Banik P, Chakraborty S, Borkar SA, Dey G, et al. Insights into molecular therapy of glioma: current challenges and next generation blueprint. *Acta Pharmacol Sin* (2017) 38:591–613. doi: 10.1038/aps.2016.167
2. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1
3. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *New Engl J Med* (2015) 372:2481–98. doi: 10.1056/NEJMoa1402121
4. Hochberg FH, Atai NA, Gonda D, Hughes MS, Mawejje B, Balaj L, et al. Glioma diagnostics and biomarkers: an ongoing challenge in the field of medicine and science. *Expert Rev Mol Diagnostics* (2014) 14:439–52. doi: 10.1586/14737159.2014.905202
5. Anderson KS, LaBaer J. The sentinel within: exploiting the immune system for cancer biomarkers. *J Proteome Res* (2005) 4:1123–33. doi: 10.1021/pr0500814
6. Chapman CJ, Thorpe AJ, Murray A, Parsy-Kowalska CB, Allen J, Stafford KM, et al. Immunobiomarkers in small cell lung cancer: potential early cancer signals. *Clin Cancer Res: Off J Am Assoc Cancer Res* (2011) 17:1474–80. doi: 10.1158/1078-0432.CCR-10-1363
7. Fonseca C, Dranoff G. Capitalizing on the immunogenicity of dying tumor cells. *Clin Cancer Res: Off J Am Assoc Cancer Res* (2008) 14:1603–8. doi: 10.1158/1078-0432.CCR-07-2245
8. Hudson ME, Pozdnyakova I, Haines K, Mor G, Snyder M. Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. *Proc Natl Acad Sci U States America* (2007) 104:17494–9. doi: 10.1073/pnas.0708572104
9. King CL, Davies DH, Felgner P, Baum E, Jain A, Randall A, et al. Biosignatures of Exposure/Transmission and Immunity. *Am J Trop Med Hyg* (2015) 93:16–27. doi: 10.4269/ajtmh.15-0037
10. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res* (2018) 46:W486–94. doi: 10.1093/nar/gky310
11. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* (2004) 3:Article3. doi: 10.2202/1544-6115.1027
12. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* (2004) 5:R80. doi: 10.1186/gb-2004-5-10-r80
13. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* (2009) 4:44–57. doi: 10.1038/nprot.2008.211
14. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* (2013) 8:1551–66. doi: 10.1038/nprot.2013.092
15. McKeon A, Pittock SJ, Lennon VA. CSF complements serum for evaluating paraneoplastic antibodies and NMO-IgG. *Neurology* (2011) 76:1108–10. doi: 10.1212/WNL.0b013e318211c379

16. Wang SS, Smiraglia DJ, Wu YZ, Ghosh S, Rader JS, Cho KR, et al. Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning. *Cancer Res* (2008) 68:2489–97. doi: 10.1158/0008-5472.CAN-07-3194
17. Demokan S, Chuang AY, Pattani KM, Sidransky D, Koch W, Califano JA. Validation of nucleolar protein 4 as a novel methylated tumor suppressor gene in head and neck cancer. *Oncol Rep* (2014) 31:1014–20. doi: 10.3892/or.2013.2927
18. Ramalho-Carvalho J, Goncalves CS, Graca I, Bidarra D, Pereira-Silva E, Salta S, et al. A multiplatform approach identifies miR-152-3p as a common epigenetically regulated onco-suppressor in prostate cancer targeting TMEM97. *Clin Epigenet* (2018) 10:40. doi: 10.1186/s13148-018-0475-2
19. Takahashi S, Shiraishi T, Miles N, Trock BJ, Kulkarni P, Getzenberg RH. Nanowire analysis of cancer-testis antigens as biomarkers of aggressive prostate cancer. *Urology* (2015) 85:704 e1–7. doi: 10.1016/j.urology.2014.12.004
20. Stangeland B, Mughal AA, Grieg Z, Sandberg CJ, Joel M, Nygard S, et al. Combined expressional analysis, bioinformatics and targeted proteomics identify new potential therapeutic targets in glioblastoma stem cells. *Oncotarget* (2015) 6:26192–215. doi: 10.18632/oncotarget.4613
21. Mains RE, Alam MR, Johnson RC, Darlington DN, Back N, Hand TA, et al. Kalirin, a multifunctional PAM COOH-terminal domain interactor protein, affects cytoskeletal organization and ACTH secretion from AtT-20 cells. *J Biol Chem* (1999) 274:2929–37. doi: 10.1074/jbc.274.5.2929
22. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. *Nature* (2007) 446:153–8. doi: 10.1038/nature05610
23. Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, et al. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* (2002) 417:967–70. doi: 10.1038/nature00769
24. Freed EF, Baserga SJ. The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis. *Nucleic Acids Res* (2010) 38:4798–806. doi: 10.1093/nar/gkq185
25. Yu B, Mitchell GA, Richter A. Nucleolar localization of cirhin, the protein mutated in North American Indian childhood cirrhosis. *Exp Cell Res* (2005) 311:218–28. doi: 10.1016/j.yexcr.2005.08.012
26. Guo F, Chen JJ, Tang WJ. CIRH1A augments the proliferation of RKO colorectal cancer cells. *Oncol Rep* (2017) 37:2375–81. doi: 10.3892/or.2017.5497
27. Yu B, Mitchell GA, Richter A. Cirhin up-regulates a canonical NF-kappaB element through strong interaction with Cirip/HIVEP1. *Exp Cell Res* (2009) 315:3086–98. doi: 10.1016/j.yexcr.2009.08.017
28. Xia Y, Shen S, Verma IM. NF-kappaB, an active player in human cancers. *Cancer Immunol Res* (2014) 2:823–30. doi: 10.1158/2326-6066.CIR-14-0112
29. Petit A, Ragu C, Soler G, Ottolenghi C, Schluth C, Radford-Weiss I, et al. Functional analysis of the NUP98-CCDC28A fusion protein. *Haematologica* (2012) 97:379–87. doi: 10.3324/haematol.2011.047969
30. Reuschenbach M, von Knebel Doeberitz M, Wentzensen N. A systematic review of humoral immune responses against tumor antigens. *Cancer Immunol Immunother: CII* (2009) 58:1535–44. doi: 10.1007/s00262-009-0733-4
31. Bernat RL, Borisy GG, Rothfield NF, Earnshaw WC. Injection of anticentromere antibodies in interphase disrupts events required for chromosome movement at mitosis. *J Cell Biol* (1990) 111:1519–33. doi: 10.1083/jcb.111.4.1519
32. Guldner HH, Szostecki C, Vosberg HP, Lakomek HJ, Penner E, Bautz FA. Scl 70 autoantibodies from scleroderma patients recognize a 95 kDa protein identified as DNA topoisomerase I. *Chromosoma* (1986) 94:132–8. doi: 10.1007/BF00286991
33. Shero JH, Bordwell B, Rothfield NF, Earnshaw WC. High titers of autoantibodies to topoisomerase I (Scl-70) in sera from scleroderma patients. *Science* (1986) 231:737–40. doi: 10.1126/science.3003910
34. Zhang JY, Zhu W, Imai H, Kiyosawa K, Chan EK, Tan EM. De-novo humoral immune responses to cancer-associated autoantigens during transition from chronic liver disease to hepatocellular carcinoma. *Clin Exp Immunol* (2001) 125:3–9. doi: 10.1046/j.1365-2249.2001.01585.x

Conflict of Interest: Author PS was employed by company Inme Oy and has no competing financial interests.

The remaining 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 © 2020 Gahoi, Syed, Choudhary, Epari, Moiyadi, Varma, Gandhi and Srivastava. 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 Oncology

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/fonc.2020.606765

KIF21B Expression in Osteosarcoma and Its Regulatory Effect on Osteosarcoma Cell Proliferation and Apoptosis Through the PI3K/AKT Pathway

Songjia Ni[†], *Jianjun Li*[†], *Sujun Qiu*, *Yingming Xie*, *Kaiqin Gong and Yang Duan*[*]

Department of Orthopedics, Zhujiang Hospital, Southern Medical University, Guangzhou, China

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Fuming Li, University of Pennsylvania, United States
Alexandre Bruni-Cardoso, University of São Paulo, Brazil

***Correspondence:**
Yang Duan
duanyang2011@163.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *15 September 2020*
Accepted: *10 December 2020*
Published: *28 January 2021*

Citation:
Ni S, Li J, Qiu S, Xie Y, Gong K and Duan Y (2021) KIF21B Expression in Osteosarcoma and Its Regulatory Effect on Osteosarcoma Cell Proliferation and Apoptosis Through the PI3K/AKT Pathway. Front. Oncol. 10:606765. doi: 10.3389/fonc.2020.606765

Osteosarcoma (OS) is the most common malignancy that occurs mainly during childhood and adolescence; however, no clear molecular or biological mechanism has been identified. In this study, we aimed to explore new biomarkers for the early diagnosis, targeted treatment, and prognostic determination of osteosarcoma. We first used bioinformatics analysis to show that KIF21B can be used as a biomarker for the diagnosis and prognosis of osteosarcoma. We then examined the expression of KIF21B in human osteosarcoma tissues and cell lines using immunohistochemistry, western blotting, and qRT-PCR. It was found that KIF21B expression was significantly upregulated in osteosarcoma tissues and cell lines. After knocking down the expression of KIF21B in the osteosarcoma cell lines 143B and U2-OS, we used cell fluorescence counting, CCK-8 assays, flow cytometry, and TUNEL staining to examine the effects of KIF21B on osteosarcoma cell proliferation and apoptosis. The results demonstrated that knocking down KIF21B in 143B and U2-OS cells could increase cell apoptosis, inhibit cell proliferation, and reduce tumor formation in nude mice. Subsequently, we used gene chips and bioinformatics to analyze the differential gene expression caused by knocking down KIF21B. The results showed that KIF21B may regulate OS cell proliferation and apoptosis by targeting the PI3K/AKT pathway. We then examined the expression of PI3K/AKT- and apoptosis-related proteins using western blotting. KIF21B knockdown inhibited the PI3K pathway, downregulated Bcl-2, and upregulated Bax. Moreover, the use of PI3K/AKT pathway agonists reversed the regulatory effect of KIF21B on the apoptosis and proliferation of 143B and U2-OS cells. In conclusion, our results indicated that KIF21B plays a key role in osteosarcoma. Low KIF21B expression might indirectly increase the apoptosis and inhibit the proliferation of osteosarcoma cells through the PI3K/AKT pathway.

Keywords: KIF21B, osteosarcoma, bioinformatics, proliferation, apoptosis, PI3K/AKT pathway

INTRODUCTION

Osteosarcoma (OS) is the most common primary malignant tumor in children and adolescents (1) and occurs in long bones, such as the proximal humerus and distal femur (2). OS is prone to metastasis and relapse, and 20–25% of patients present with lung metastases at the time of diagnosis (3). The current treatment for OS involves surgery and adjuvant chemotherapy, and the prognosis is poor, with 5-year survival rates below 20% (4).

KIF21B belongs to the kinesin superfamily of proteins (5). Human kinesins are divided into 14 families that consist of 45 different kinesin proteins and are involved in a series of cellular processes, such as mitosis, motility, organelle transport, and tumor development (6). Studies have confirmed that KIF21B is expressed in many types of cells, including neurons (7, 8). However, there are few studies on the relationship between the KIF21B protein and human tumors. When preparing this manuscript, we identified one study that first observed a relationship between KIF21B and hepatocellular carcinoma (9). The function and mechanisms of KIF21B in cancer, particularly in OS, and its effect on prognosis have not yet been widely investigated and remain unknown.

The PI3K/Akt signaling pathway is composed of a conserved family of signal transducing enzymes that are involved in a wide range of physiological processes (10, 11). The PI3K/AKT signaling pathway is also one of the most commonly dysregulated pathways in cancer (12). Therefore, the PI3K/AKT pathway has become the main focus of drug development for the treatment of cancer (13). Moreover, the PI3K/AKT signaling pathway also plays an important role in the regulation of various kinesin proteins (14, 15).

In this study, we first identified KIF21B as a potential new biomarker of OS through bioinformatics analysis. To explore the impact of KIF21B on OS, we examined the expression of KIF21B in osteosarcoma tissues and cell lines. Then, we evaluated the effects of KIF21B on the proliferation and apoptosis of 143B and U2-OS osteosarcoma cells. We further analyzed the differential gene expression caused by knocking down KIF21B and predicted and verified the potential downstream target genes of KIF21B. The purpose of this study was to investigate the role of KIF21B in the progression of OS and to provide ideas for new treatment options for OS.

MATERIALS AND METHODS

Data Processing

The OS-related chip data were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo), including GSE12865 (12 osteosarcoma cells and 2 normal human osteoblastic cells). The differentially expressed mRNAs were analyzed with an adjusted P value <0.05 and |Log2FC| >0.25 as the threshold using the limma package in R software. The dysregulated mRNAs were used for future analysis, and the expression levels of the mRNAs were visualized by volcano plots using the ggplot2 package in R. In addition, the RNA-seq data and clinical features of OS patients were collected from the TCGA database (TARGET-OS, N = 85). The characteristics of the patients included in this study are summarized in **Table 1**.

Protein-Protein Interaction Network Analysis

The interactive relationships among the dysregulated genes were analyzed using a protein-protein interaction (PPI) network. The gene information was retrieved from the Search Tool for the Retrieval of Interacting Genes (STRING) database (version 11.0, https://string-db.org/). The PPI network was constructed and visualized by using Cytoscape 3.7.2 software. The ClusterONE algorithm was used to mine the hub modules in the PPI network. The threshold parameters were set to minimum size = 50 and density threshold = 0.5. The nodes with a high degree of connectivity (top 15, degree ≥9) were considered hub genes involved in important biological functions.

Functional Enrichment and GO Semantic Similarity

Enrichment analysis of the dysregulated genes was used to investigate the biochemical pathways involved in carcinogenesis and tumor progression. The GO functional enrichment analysis and KEGG pathway analysis were performed using the clusterProfiler package in R. We set the criterion that the P value of the GO terms and KEGG analysis was less than 0.05. The top 15 enriched GO terms and KEGG pathways were visualized using the ggplot2 package in R. The functional similarity between genes was measured by semantic similarity with GO terms [17344234]. The semantic similarity computation among genes was performed using the GOSemSim R package.

Receiver Operating Characteristic Curve Analysis

Receiver operating characteristic (ROC) curve analysis was used to evaluate the diagnostic value of the target mRNAs in patients with OS by calculating the sensitivity and specificity. The area under the curve (AUC) of the ROC curve was used to assess the

TABLE 1 | The information of clinical features.

Characteristics	Number of cases	Percentages
Age		
<20	74	87.06
>20	11	12.94
Gender		
Female	37	43.53
Male	48	56.47
Race		
Asian	6	7.06
Black or African American	7	8.24
White	51	60.00
Others	21	24.70

discrimination performance. An AUC above 0.7 was considered satisfactory diagnostic accuracy [29901123].

Survival, Univariable, and Multivariable Analyses

Kaplan-Meier survival analysis was performed by using the survival R package to predict the overall survival of the OS patients. The stratified log-rank test was used to compare the differences in overall survival time between the two groups. The Cox proportional hazards regression model was used to explore the univariable and multivariable hazard ratios (HRs) and screen the OS prognostic factors [31216997]. The univariable Cox proportional hazards regression model was used to select the independent clinicopathological factors. Subsequently, the multivariable Cox proportional hazards regression model was implemented to explore the potential prognostic biomarkers of OS. The patients with OS were divided into high-risk and low-risk groups according to the risk score in the hazards model. The time-dependent ROC curve was used to assess the risk prediction rate of the target genes between the high-risk and low-risk groups. All the statistical analyses were performed using R and SPSS 22.0 software. A P value less than 0.05 was considered statistically significant.

Specimens

A total of 17 pairs of OS tissues and corresponding adjacent tissues were collected from OS patients in the Orthopedics Department of Zhujiang Hospital of Southern Medical University. All the pathological specimens were diagnosed by postoperative pathology and derived from newly formed cancerous tissues with complete clinical data. The patients did not receive preoperative chemoradiotherapy. The patients' informed consent was acquired for all the procedures related to their tissues, and the study was approved by the Ethics Committee of the Zhujiang Hospital of Southern Medical University.

Cell Culture

The human osteoblast cell line hFOB1.19 and the human OS cell lines 143B, U2-OS and MG63 were purchased from the American Type Culture Collection (ATCC, Rockville, MD, USA). The cells were cultured in RPMI 1640 medium (HyClone Technologies, Logan, USA) containing 10% fetal bovine serum (FBS, Thermo Fisher Scientific, MA, USA) and 100 U/ml anti-bis (Thermo Fisher Scientific) in 5% CO2 at 37°C.

Immunohistochemistry

Human tissues were fixed in 10% formalin (Solarbio, Beijing, China) and embedded in paraffin. Five-micrometer sections were dewaxed with dimethylbenzene (Macklin, Shanghai, China) and dehydrated in a gradient ethanol series, followed by incubation with citrate buffer for antigen recovery. Hydrogen peroxide (Macklin, 3%) was used to block the endogenous peroxidase activity. The sections were incubated with primary antibodies against KIF21B (1:500; Thermo Fisher) overnight at 4°C. The sections were then stained using biotinylated secondary antibodies at 37°C for 20 min and exposed using DAB. The images were captured by microscopy (Leica, Wetzlar, Hesse-Darmstadt, Germany). Three fields were randomly selected, and the average percentage of brown or dark yellow particles in the cytoplasm was counted.

Western Blotting

The total protein was extracted using a whole cell lysis assay (KeyGEN Biotech, Nanjing, Jiangsu Province, China). Eighty micrograms of sample protein was subjected to SDS-PAGE (KeyGEN Biotech) and transferred to PVDF membranes (Millipore, MA, USA). The membranes were blocked and probed with the indicated primary antibodies at 4°C for 12 h. The membranes were then incubated with the indicated HRP-conjugated secondary antibodies at room temperature for 2 h, and the expression of the target proteins was detected by ECL (KeyGEN Biotech). The following antibodies were used: KIF21B (1:3,000; Thermo Fisher), PI3K (1:3,000; Thermo Fisher), p-PI3K (1:1,000; Thermo Fisher), AKT (1:3,000; Thermo Fisher), p-AKT (1:1,000; Thermo Fisher), Bcl-2 (1:2,000; Cell Signaling Technology, Boston, MA, USA), BAX (1:1,000; Cell Signaling Technology), and Actin (1:1,000; Beyotime Biotechnology, Shanghai, China).

qRT-PCR

The total RNA was extracted and synthesized into cDNA according to the manufacturer's protocol (TaKaRa Bio, Dalian, Liaoning Province, China). qRT-PCR was performed on a LightCycler 96 (Roche Life Sciences, Switzerland, Basel) using Real-Time PCR Mix (Vazyme Biotech, Nanjing, Jiangsu Province, China). Gene expression relative to GAPDH expression was assessed using the 2-ΔΔCt method. Independent experiments were conducted in triplicate (KIF21B upstream sequence: 5'-GGATGCCACAGATGAGTT-3', downstream sequence: 5'-TGTCCCGTAACCAAGTTC-3. GAPDH upstream sequence: 5′-ATAGCACAGCCTGGATAGCAACGTAC-3′, downstream sequence: 5′-CACCTTCTACAATGAGCTGCGTGTG-3′.)

Lentivirus Infection

A short hairpin RNA (shRNA) sequence specific for human KIF21B (GGAGCTGATGGAGTATAAG) and a negative control sequence (TTCTCCGAACGTGTCACGT) were constructed and confirmed. Lentivirus with and without fluorescence were provided and titered by GeneChem (Shanghai, China). The cells in the shKIF21B group were transduced with shRNA lentivirus with a multiplicity of infection (MOI) of 20. The shCtrl group was transduced with a negative control sequence, and the blank control group was not treated. Seventy-two hours after transduction, the cells were continuously cultured for 1 week using complete medium containing puromycin (2 μg/ml, Beyotime Biotechnology).

Cell Growth Assessments

Virus-infected cells exhibiting green fluorescence were counted on a Celigo instrument (Nexcelom, Lawrence, MA, USA). After

3–5 days of continuous measurements, cell growth curves were plotted to reflect cell growth.

Cell Counting Kit-8

Cells from each group were seeded into a 96-well plate at a density of 2,000 cells/well, and 10 μl CCK-8 solution (Dojindo, Kumamoto, Japan) was added to each well every 24 h. After incubation at 37°C for 2 h, the absorbance value at 450 nm was detected using a microplate reader (Bio-Rad, Hercules, USA). Triplicate wells were used for all the specimens in each test.

Cell Apoptosis Analysis by Flow Cytometry

The Annexin FITC/PI Apoptosis Detection Kit (KeyGEN Biotech) was used following the manufacturer's protocols. A total of 1 × 106 cells were resuspended in a mixture of 100 μl of binding buffer, 10 μl of FITC Annexin V, and 10 μl of PI solution and then incubated in the dark for 15 min at room temperature. Then, cell apoptosis was measured using a flow cytometer (BD Biosciences, NJ, USA). The results were analyzed using FlowJo software. The experiment was repeated in triplicate.

TUNEL Assay Kit

Adherent cells were sequentially fixed with 4% paraformaldehyde (Beyotime Biotechnology) for 30 min and incubated with Triton X-100 (Beyotime Biotechnology) for 5 min at room temperature. Then, the TUNEL test solution was prepared according to the operating instructions (Beyotime Biotechnology). The cells were added to 50 μl TUNEL detection solution and incubated at 37°C for 60 min in the dark, and the nuclei were stained with DAPI (Beyotime Biotechnology). The nuclei and TUNEL-positive cells were then observed under a fluorescence microscope (Nikon, Tokyo, Japan), and the positive rate (TUNEL/DAPI) of five random fields was calculated.

Grouping and Treatment of Nude Mice

Nine BALB/c nude mice (6 weeks, approximately 18 g, Animal Center of Southern Medical University) were used for the tumor formation experiments. A total of 3 × 106 143B cells were subcutaneously injected into the right forelegs of the nude mice. A Vernier caliper was used to measure the longest and shortest diameters of the transplanted tumor every 5 days, the volume of the subcutaneous transplanted tumor was calculated ($V = L \times W2 \times 0.5$), and the tumor growth curve was drawn. One month after the injection, the mice were anesthetized by using carbon dioxide. The tumor tissue was then removed, weighed, and imaged. The animal experiment was performed in strict accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Gene Microarray Hybridization, Scanning, and Analysis

Hybridization of nucleic acid probes and gene expression profiling chips were performed with a gene chip hybridization kit (Thermo Fisher) for 16 h. Elution and staining were performed using Onfly Fluid Workstation 450. A Gene Chip Scanner 3000 7G scan (Affymetrix, Thermo Fisher) was used to generate images. Chip scan image 5.0 software was used for digital processing and analysis.

Statistical Analysis

All the data were analyzed using SPSS 22.0 software. The qualitative data are presented as counts (%), and the quantitative data are expressed as the mean ± standard deviation. The measurement data that conformed to a normal distribution were assessed *via* T-tests. Mann-Whitney rank sum tests were used to analyze the data that did not conform to a normal distribution. Three groups of data were analyzed through one-way analysis of variance. Differences were statistically significant at the $P < 0.05$ level.

RESULTS

Comprehensive Analysis Identified KIF21B as a Potential Diagnostic and Prognostic Gene in OS

First, principal component analysis (PCA) was performed to identify abnormal samples (**Supplementary Figure 1**). Then, the differentially expressed genes were analyzed by using the limma package in R software. A total of 10,446 differentially expressed genes (DEGs) were identified between the osteosarcoma cells and normal human osteoblastic cells with an adjusted P value <0.05 and $|Log2FC| > 0.25$ as the threshold. There were 4,875 upregulated and 5,571 downregulated mRNAs (**Figure 1A**). To screen and predict the possible diagnostic mRNAs in OS, the diagnostic values were analyzed using ROC curves. A total of 10,441 DEGs were used in the subsequent studies with AUC > 0.7 as the cutoff criterion (**Supplementary Table 1**). To select novel genes for predicting the overall survival of patients with OS, survival analysis was performed by the Kaplan-Meier method with the log-rank test. The results showed that 1,561 novel genes were associated with the overall survival of OS patients, with a P value <0.05 (**Supplementary Table 2**). To identify the potential prognostic indicators of OS, univariable Cox regression analysis was performed to identify independent risk factors for the prognosis of OS. Here, we identified 681 DEGs that were significantly associated with cancer-specific survival with a P value <0.01 as the threshold (**Supplementary Table 3**). Subsequently, 496 potential target genes were obtained as effective diagnostic and prognostic genes (**Figure 1B**).

To identify the interactions between these 496 target genes, a PPI network was constructed by using the STRING database, with scores >950 as the threshold. The results showed that 100 genes were related to each other in the PPI network (**Figure 1C**). Moreover, the ClusterONE algorithm was particularly useful for detecting protein complexes in the PPI network. A total of 11 PPI modules were identified, with a P value <0.05 (**Figure 1D**). The hub genes were considered to play important roles in the highly intersected network. Fifteen hub genes were identified by network clustering analysis with degree ≥9 as the cutoff criteria (**Figure 1E**).

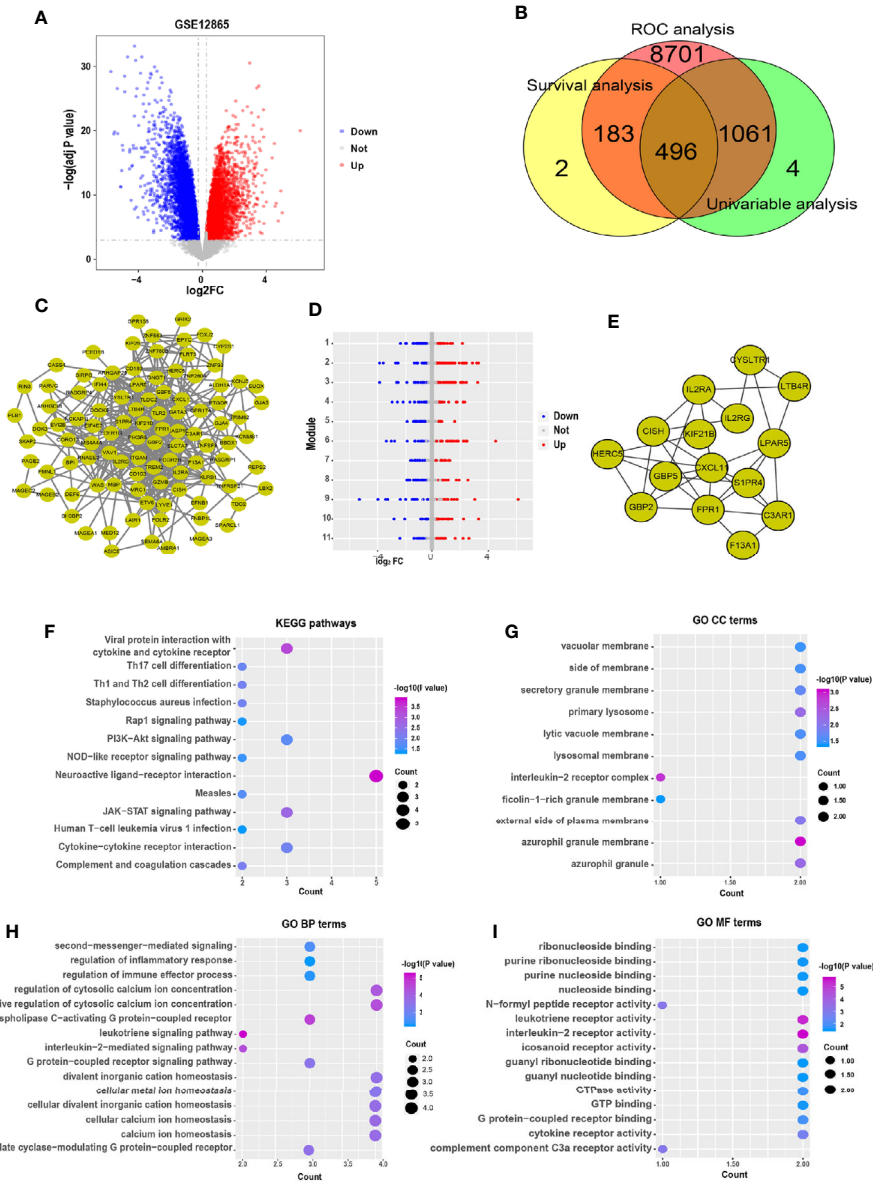


FIGURE 1 | Comprehensive expression analysis for the identification of hub genes associated with the progression of osteosarcoma. **(A)** The volcano plot of significant differentially expressed genes. The red dots represent the upregulated genes, and the blue dots represent the downregulated genes. **(B)** Venn diagram representing the number of dysregulated genes in the ROC, survival and univariable analyses. **(C)** A PPI network of the dysregulated genes was constructed. **(D)** The modules were identified from the PPI network using the ClusterONE algorithm, containing 11 modules. The red dots represent the upregulated genes, the blue dots represent the downregulated genes, and the gray dots represent the genes with no significant difference in expression. **(E)** Interaction network of hub genes. **(F)** Significant pathways in the KEGG pathway analysis with $P < 0.05$. **(G)** Eleven GO terms were enriched for cellular components. **(H)** The top 15 enriched biological process GO terms of the dysregulated genes. **(I)** The top 15 enriched molecular functions GO terms.

To explore the biological functions of these hub genes, KEGG pathway and GO term enrichment analyses were performed using the clusterProfiler package in R. The results of the KEGG pathway analysis showed that these dysregulated genes were enriched in cancer-associated pathways, including the JAK-STAT, PI3K-Akt, NOD-like receptor, and Rap1 signaling pathways (**Figure 1F**). In addition, our results showed that 11 GO terms were enriched for cellular components (CC), 125 GO terms were enriched for biological processes (BP), and 31 GO terms were enriched for molecular functions (MF), with $P < 0.05$ as the threshold. The CC GO terms included those associated with secretory granule membrane and vacuolar membrane (**Figure 1G**), and the BP GO terms mainly included those associated with calcium ion homeostasis, second messenger

and regulation of immune effector process (**Figure 1H**). The MF GO terms were associated with G protein-coupled receptor binding and nucleoside binding (**Figure 1I**).

In addition, multivariable Cox hazard regression analysis was performed to examine the independent risk factors. We constructed 15 Cox models using the hub genes and clinical features, including age, sex, and race. The results showed that 10 Cox models were identified as independent factors influencing survival time (P value < 0.01; Wald test), including KIF21B (HR = 0.094, 95% CI: 0.018–0.501, P = 0.006), LTB4R (HR = 0.047, 95% CI: 0.007–0.336, P = 0.002), CISH (HR = 0.048, 95% CI: 0.007–0.316, P = 0.002), GBP2 (HR = 0.135, 95% CI: 0.035–0.517, P = 0.003), CYSLTR1 (HR = 0.147, 95% CI: 0.040–0.533, P = 0.004), F13A1 (HR = 0.176, 95% CI: 0.070–0.445, P < 0.0001), GBP5 (HR = 0.296, 95% CI: 0.136–0.642, P = 0.002), FPR1 (HR = 0.358, 95% CI: 0.182–0.701, P < 0.003), IL2RA (HR = 0.474, 95% CI: 0.294–0.767, P = 0.002), and S1PR4 (HR = 0.505, 95% CI: 0.308–0.827, P = 0.007). Visualizations of the top five results are shown in **Figure 2A**. Next, the key gene expression-based survival risk score of the 85 patients was calculated. All the samples were divided into two groups, the high- and low-risk groups, according to the median risk score. The results showed that the high-risk score was associated with poorer overall survival (**Figure 2B**). Furthermore, a time-dependent ROC curve was used to assess the power of the prognostic factors. The AUC values for the prognostic models ranged from 0.710 to 0.792 at 3 years of overall survival (**Figure 2C**). Moreover, to identify novel biomarkers for further study, GO semantic similarity was used to quantify the functional similarity between 10 key genes. The results showed that KIF21B was considered to be protein with the most significantly strengthened relationship with its partners (**Figure 2D**). The expression of KIF21B was analyzed in the high- and low-risk patients with OS. The results showed that KIF21B was upregulated in the high-risk patients (**Figure 2E**). Taken together, these findings provide evidence that KIF21B is a diagnostic and prognostic biomarker and potential therapeutic target in patients with OS.

Expression of KIF21B in Human OS Tissues and Cell Lines

To determine and assess the expression level of KIF21B in OS, the expression of KIF21B in osteosarcoma cells and normal human osteoblastic cells from the GEO database (GSE12865) was analyzed (**Figure 2F**). The results showed that KIF21B was expressed at high levels in patients with OS (adjusted P value = 0.0448). To explore the effect of KIF21B on the development of OS, we then compared the expression levels of KIF21B between tumor tissue, junction tissue and bone tissue. Immunohistochemical staining was performed to detect KIF21B expression. The results showed that the expression of KIF21B in the tumor tissues was significantly higher than that in the junction and bone tissues (**Figures 3A, B**). High KIF21B expression at the tumor site was also confirmed by western blotting analysis (**Figures 3C, D**) and qRT-PCR (**Figure 3E**). Then, qRT-PCR and western blotting were used to explore the expression of KIF21B in OS cell lines. The results demonstrated that the expression of KIF21B in the OS cell lines U2-OS, 143B, and MG63 was higher than that in the Hfob1.19 cell line (**Figures 3F–H**). These results demonstrated that KIF21B was typically overexpressed in OS, suggesting that it might be considered an oncogene.

Effects of KIF21B Silencing on OS Cell Proliferation and Apoptosis

To investigate the potential role of KIF21B in OS cell function, the expression of KIF21B was assessed 48 h after the transduction of 143B and U2-OS cells with shRNA specific for KIF21B. The KIF21B expression level in the shKIF21B group was significantly downregulated compared with that in the shCtrl group and blank control group (**Figures 4A, B**). Following transduction with the KIF21B lentivirus, the cells were assessed for 4 days, and the growth curves were plotted. The proliferation rate of the 143B cells in the shKIF21B group was significantly reduced compared with that in the control groups (**Figures 4C, D**). The CCK-8 results also demonstrated a significantly reduced absorbance value in the shKIF21B group (**Figure 4E**). In addition, a similar trend was also observed in U2-OS cells as well (**Figures 4F–H**). Moreover, the flow cytometry (**Figures 5A, B**) and TUNEL staining (**Figures 5C–F**) results showed an increased apoptosis rate in the cells transduced with shKIF21B. To confirm these findings *in vivo*, 143B cells were implanted into nude mice (**Figures 6A, B**), and the effects of KIF21B were assessed after 30 days of treatment. We measured the tumor size and weight (**Figures 6C, D**). Compared with the control, KIF21B reduced the tumor size and weight. These results indicated that low KIF21B expression could enhance the apoptosis and suppress the proliferation of OS cells.

Screening of Differentially Expressed Genes and Analysis of Bioinformatics

To analyze the possible downstream targets of KIF21B, we used a human genome-wide expression chip, which contains 49,395 sets of probes, for further research. RNA probes from each treatment group were hybridized with the chip to analyze the changes in the gene expression profiles of osteosarcoma cells before and after KIF21B knockdown. The differentially expressed RNAs between the KIF21B-silenced OS cells and normal OS cells were analyzed to obtain heatmaps, scatter plots, and volcano plots (**Figures 6E–G**), the dots of the scatter plot in the parallel green solid lines represent the genes that are not significantly different, and the red and green dots represent the genes that are relatively upregulated or downregulated in the shKIF21B group, respectively. The volcanic map was plotted by fold change and FDR between the two groups of samples. The red points are the significantly differentially expressed genes screened with | fold change | ≥1.5 and FDR <0.05 as the standard, and the gray points are the other genes with no significant differences. The results showed that compared with the control, KIF21B gene silencing resulted in a decrease in the expression of 467 genes and an increase in the expression of 382 genes. The top 50 most significant differentially expressed genes are shown in **Supplementary Tables 4**, **5**. Then, the

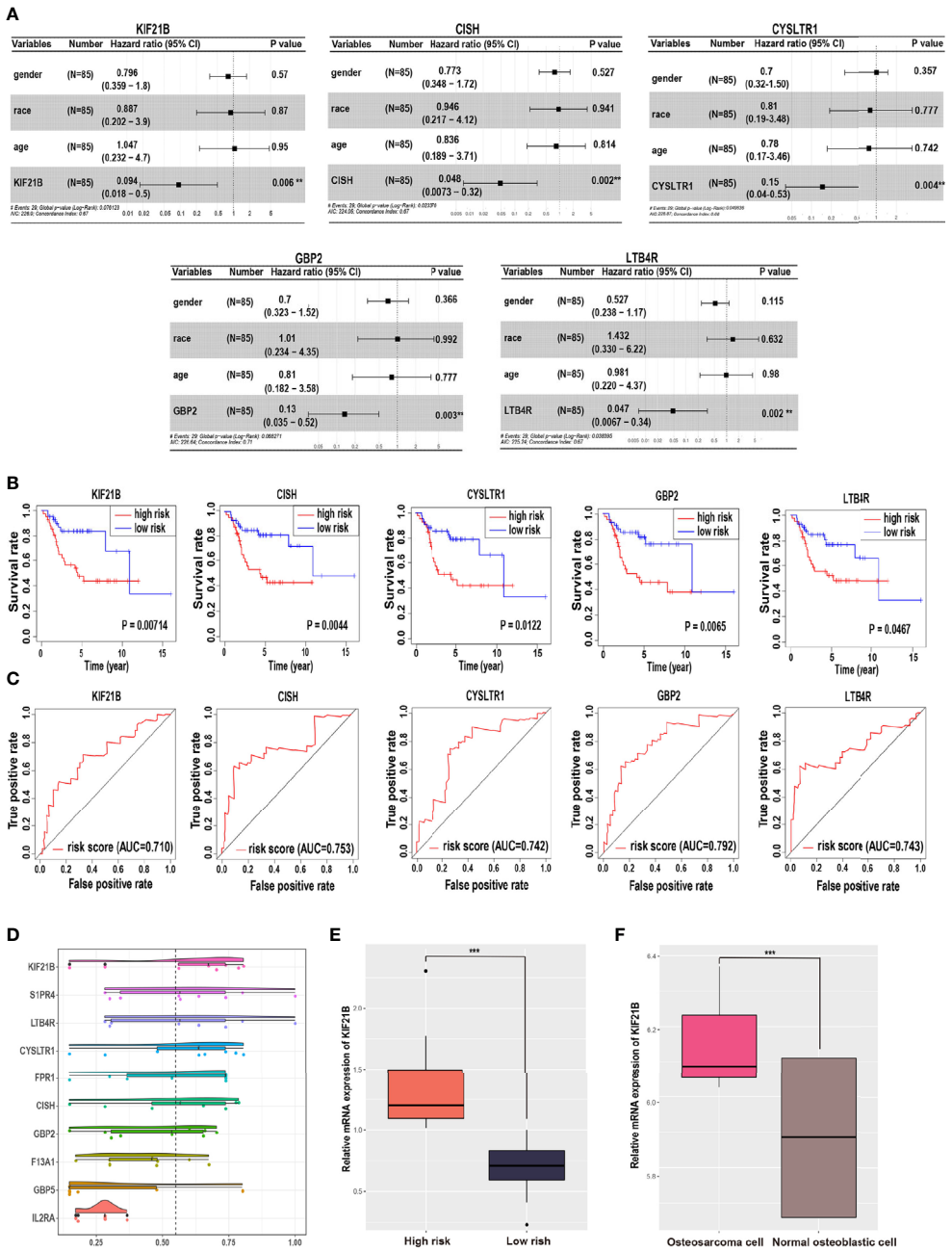


FIGURE 2 | Identification of potential prognostic biomarkers for predicting survival in patients with osteosarcoma. **(A)** Forest plot indicating the hazard ratios for overall survival using a multivariable Cox regression model. **(B)** Kaplan-Meier curves of overall survival in patients with osteosarcoma. The red lines represent high-risk group samples, and the blue lines represent low-risk group samples. **(C)** The time-dependent receiver operating characteristic curve of the prognostic signature in osteosarcoma. **(D)** Functional semantic similarity between 10 hub genes. Boxplot of KIF21B mRNA expression in the high- and low-risk groups **(E)** and in osteosarcoma cells and normal human osteoblastic cells **(F)**. *** means $P < 0.05$.

R software clusterProfiler package was used to analyze 849 differentially expressed RNAs. GO pathway analysis (**Figures 6H–J**) showed the top 20 results that were highly enriched. We found that the most enriched BP, CC, and MF terms were cell adhesion, endoplasmic reticulum membrane and protein binding, respectively. KEGG pathway analysis (**Figure 6K**) showed the top 20 pathways that were highly enriched. These pathways were mainly involved in the PI3K/Akt signaling pathway, focal adhesion, dopaminergic synapses, pathways in cancer, and MAPK signaling pathway.

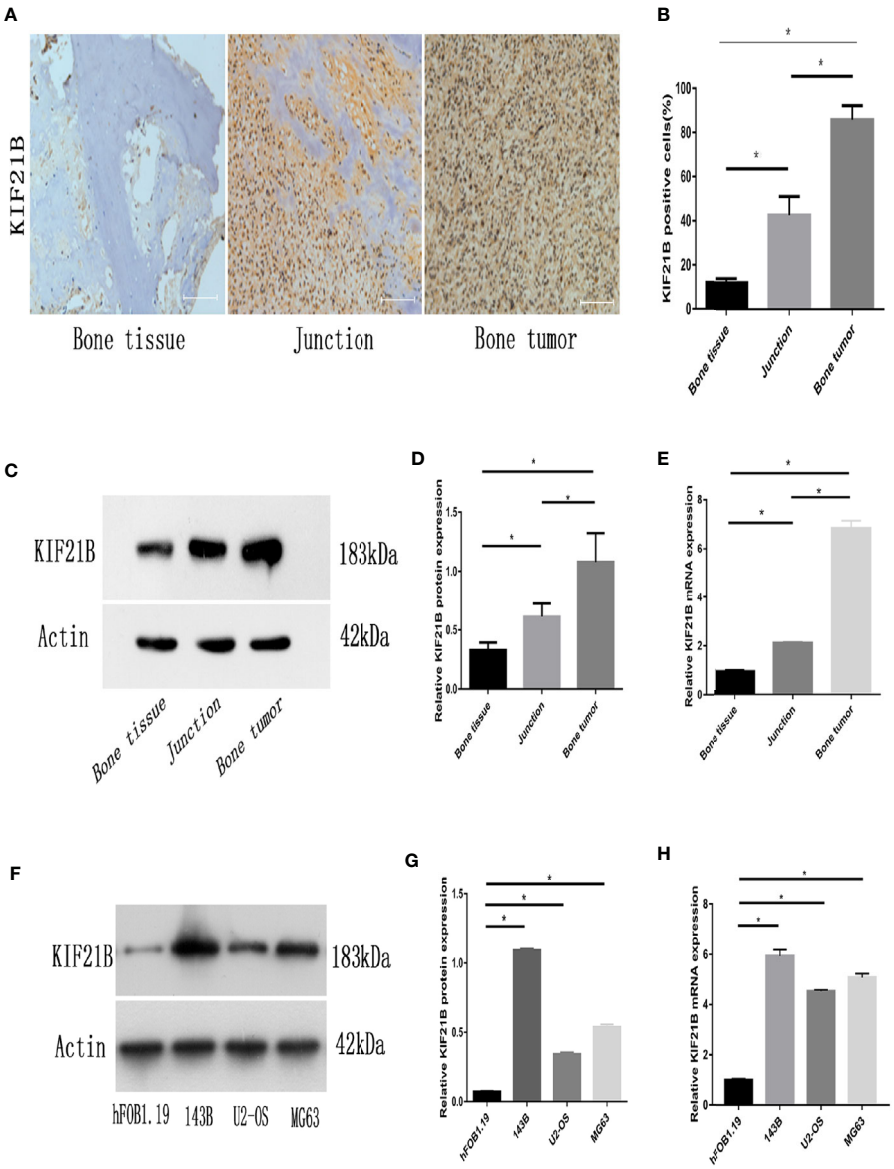


FIGURE 3 | Expression of KIF21B in human OS tissues and OS cell lines. **(A, B)** Representative immunohistochemical images (magnification, 200×) of KIF21B-positive cells show the expression of KIF21B in human osteosarcoma tissue. **(C, D)** The protein expression levels of KIF21B in the human osteosarcoma samples were detected by western blot analysis. **(E)** qRT-PCR was used to measure the mRNA expression levels of KIF21B in the human osteosarcoma samples. **(F, G)** The protein expression levels of KIF21B in human osteoblast cells hFOB1.19 and human OS cell lines 143B, U2-OS, and MG63 were detected by western blot analysis. **(H)** qRT-PCR was used to measure the mRNA expression levels of KIF21B in the hFOB1.19 and OS cell lines. Scale bar = 200 μm, *P < 0.05.

Silencing of KIF21B Affects Osteosarcoma Cell Function Through the PI3K/AKT Pathway

Functional enrichment analysis of the differentially expressed genes suggested that the PI3K/AKT signaling pathway might play an important role in the KIF21B-mediated regulation of OS cells. Since the PI3K/AKT pathway is widely involved in tumor development, we further studied the changes in the key proteins of the PI3K/AKT pathway after silencing KIF21B. The western blot analysis results suggested that the expression levels of PI3K, p-PI3K, AKT, and p-AKT were significantly downregulated after silencing KIF21B. The expression levels of Bax were upregulated, and the expression levels of Bcl-2 were downregulated (**Figure 7A**). Thus, our study demonstrated that factors involved in the PI3K/ATK signaling pathway were significantly downregulated in the shKIF21B group. In addition, based on the results of silencing KIF21B, we treated the cells with the PI3K/AKT signaling pathway agonist IGF-1, and the results showed that

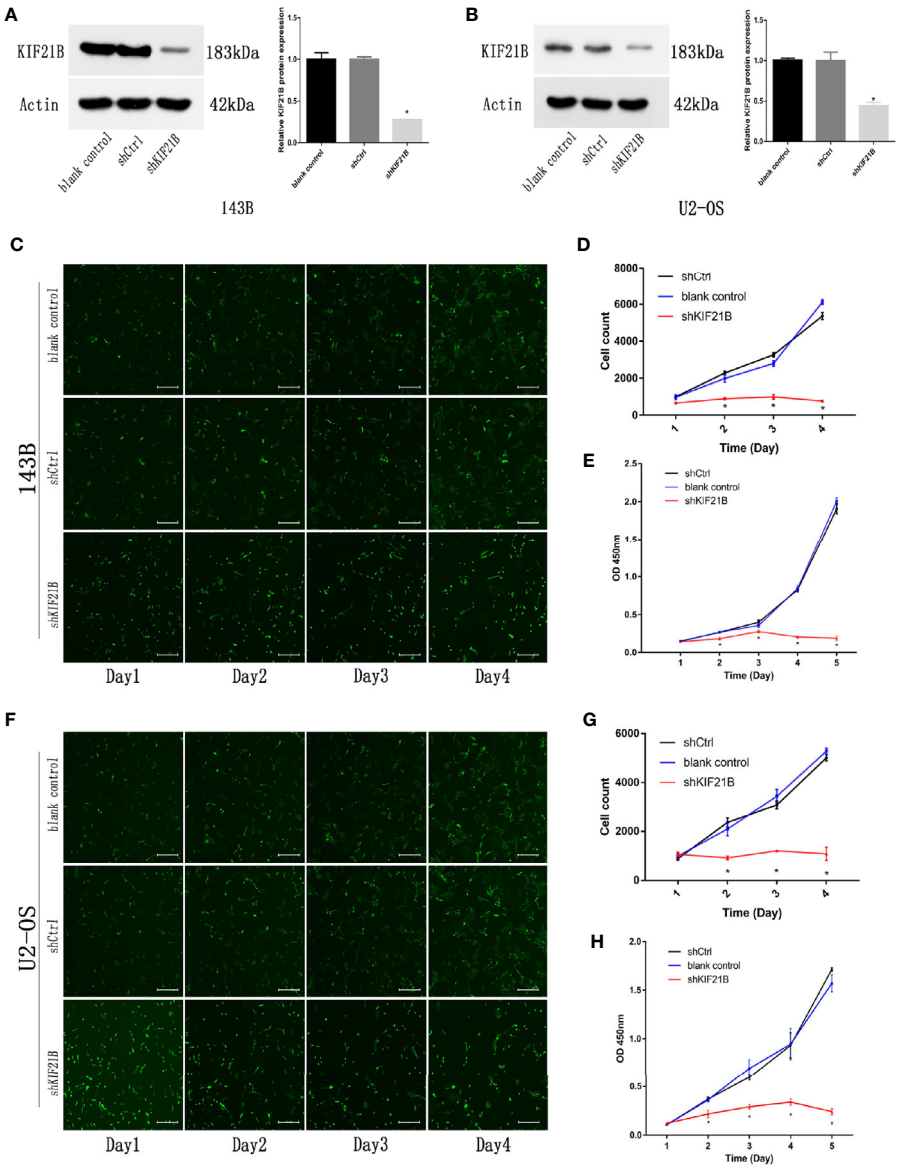


FIGURE 4 | Effects of KIF21B silencing on cell proliferation. **(A, B)** Western blotting was used to measure the KIF21B protein expression in 143B and U2-OS cells following the transfection of KIF21B shRNA. The Celigo cell count and CCK-8 results showing the proliferation of 143B **(C–E)** and U2-OS **(F–H)** cells in each group. Scale bar = 200 μm, *P < 0.05.

the inhibition of proliferation and increase in apoptosis caused by silencing KIF21B were reversed by IGF-1 (**Figures 7** and **8**). These results confirmed that the regulatory effect of KIF21B on 143B and U2-OS cell proliferation and apoptosis might be associated with the PI3K/ATK signaling pathway.

DISCUSSION

OS is the most common primary malignant bone tumor, with a particular prevalence in adolescents and children (annual incidence of ~3.1 million) (16, 17). Current treatments include a combination of neoadjuvant chemotherapy and physical rescue. Although this method preserves joint function in the extremities, OS metastasis and recurrence remain common. OS cells are highly genetically unstable and histologically heterogeneous, and their signaling is complex (18). Exploring the molecular details of OS development can improve OS diagnostics and therapeutics.

In this study, we first conducted a series of bioinformatics analyses, and the comprehensive analysis determined that KIF21B is a potential new biomarker of OS. Polymorphisms in the KIF21B gene are associated with multiple sclerosis, ankylosing

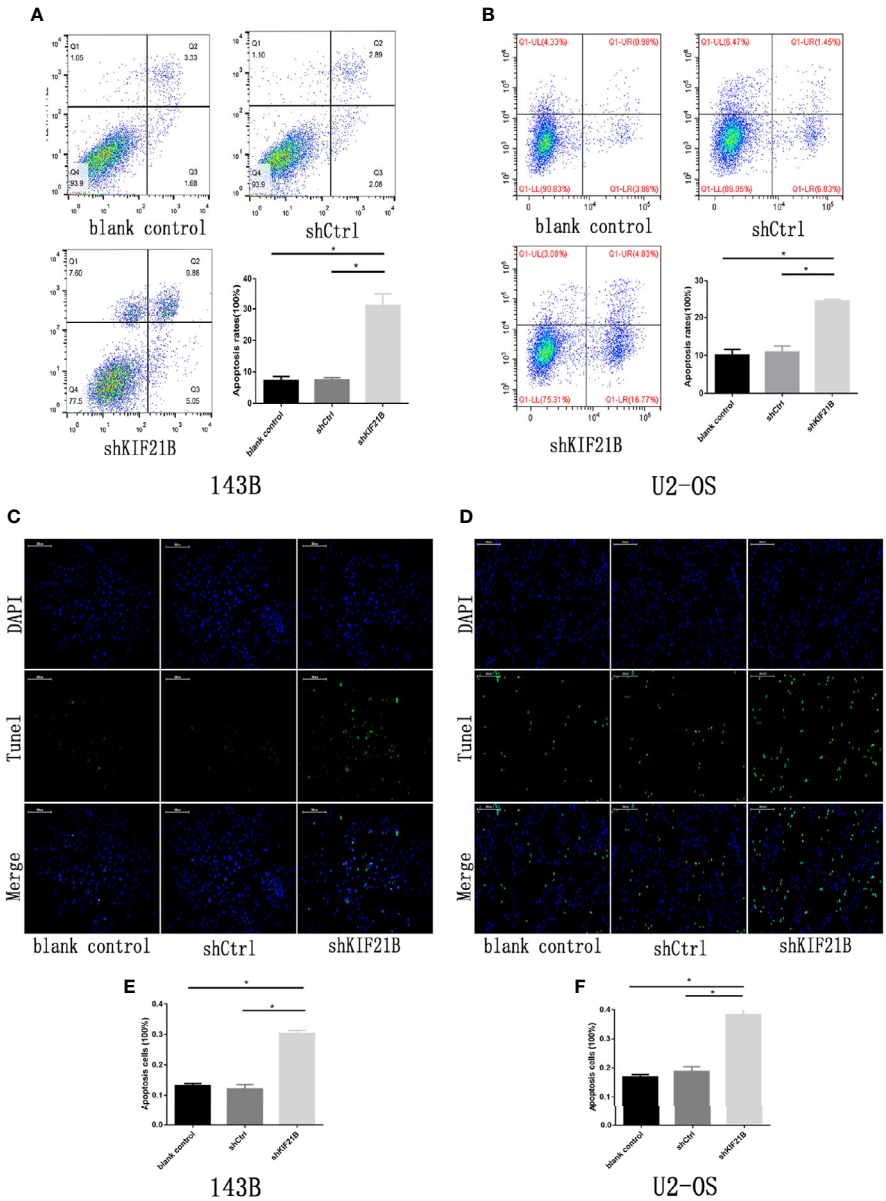


FIGURE 5 | Effects of KIF21B silencing on cell apoptosis. **(A, B)** Flow cytometry results showing the apoptosis rate of 143B and U2-OS cells in each group. **(C–F)** TUNEL assays results showing the apoptosis rate of 143B and U2-OS cells in each group. Scale bar = 200 μm, *P < 0.05.

spondylitis, Crohn's disease, and ulcerative colitis (19–21). Increased expression of KIF21B accelerates the progression of neurodegenerative diseases, such as Alzheimer's disease and multiple sclerosis (22). In addition, microrepetitions carrying the KIF21B locus are associated with neurodevelopmental and psychological abnormalities (23). Basically no research has been conducted on the role of KIF21B in human tumors. On the other hand, the abnormal expression of multiple kinesin proteins in tumor cells has been observed (24). High expression of KIF5B is observed in breast (25) and skin cancer (26), and the overexpression of KIF14 promotes the development of retinoblastoma (27), lung cancer (28), and breast cancer (29). KIF7 and Eg5/KIF11 are associated with the occurrence and metastasis of various tumors (30). The expression of KIF21B in bone tumors and its related mechanisms have not been reported. Therefore, we wanted to determine whether KIF21B plays a role in the development of osteosarcoma. Combined with the results of the bioinformatics analysis, we detected the expression of KIF21B in OS tissues and cell lines. The results demonstrated that the expression of KIF21B is significantly upregulated. We then selected the OS cell lines 143B and U2-OS. Following KIF21B silencing, we found that cell proliferation was inhibited and cell apoptosis was increased,

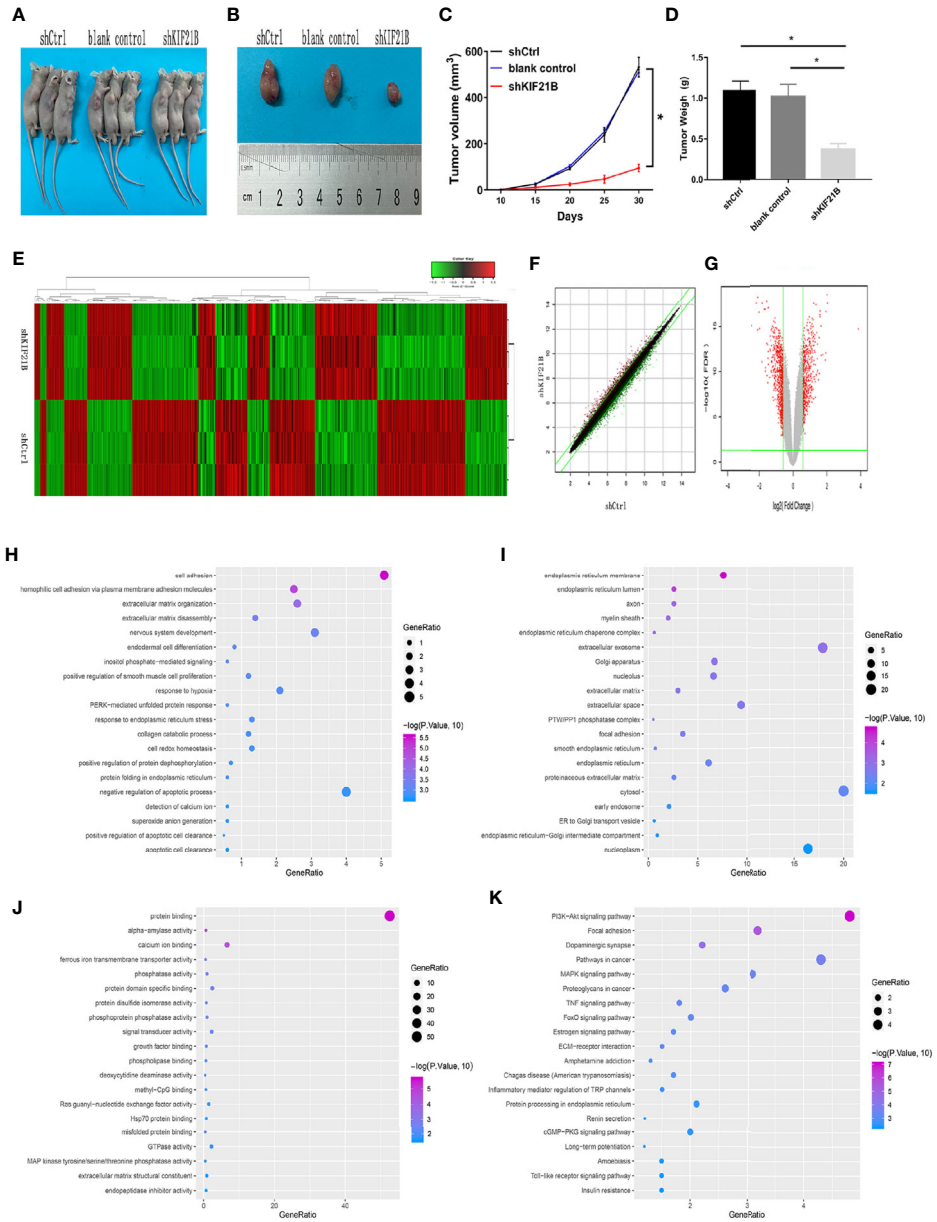


FIGURE 6 | Effects of KIF21B silencing on tumor formation in nude mice and identification of differentially expressed genes. **(A, B)** Representative photo of nude mouse tumors. The tumor size **(C)** and weight **(D)** were measured after different treatments. **(E)** Heatmap of differentially expressed genes between the shCtrl group and shKIF21B group; The scatter plot **(F)** and volcanic map **(G)** of differentially expressed genes. **(H–J)** GO annotation showed the top 20 results of biological process (BP), cellular component (CC), and molecular function (MF) that were highly enriched GO annotations. **(K)** KEGG pathway analysis showed the top 20 pathways that were highly enriched. *P < 0.05.

which indicate that KIF21B may act as an oncogene in osteosarcoma.

Under normal conditions, kinesin forms complexes with dynein molecules, which are involved in the transport of intracellular vesicles and organelles. However, the overexpression of kinesin proteins can generate additional external forces, leading to spindle collapse and monopolar spindle formation, which in turn lead to an uneven distribution of genetic material during the later stages, ultimately leading to aneuploidy. The genetic material acquired or lost in non-euploid cells is considered to be an factor that initiates the malignant progression of cancer (31). To further study the mechanism by which KIF21B participates in the development of osteosarcoma, we used a gene chip to identify differentially expressed genes before and after KIF21B silencing and performed bioinformatics analysis. According to the results, the differentially expressed genes caused by KIF21B silencing are mainly enriched in

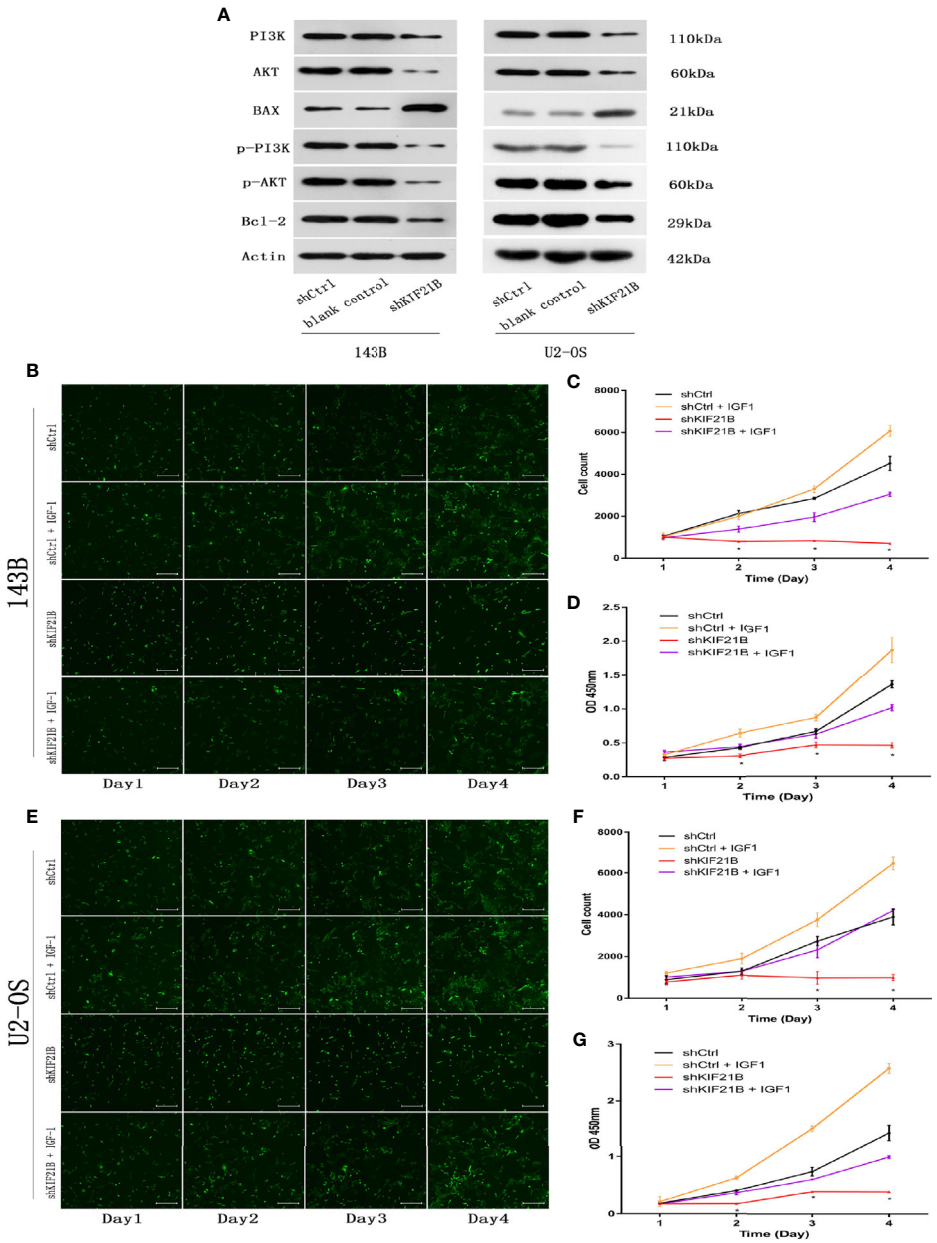


FIGURE 7 | Validation of downstream target pathways and effects of IGF-1 on cell proliferation and apoptosis after silencing KIF21B. **(A)** The western blots show the PI3K/AKT, Bax, and Bcl2 results in the shKIF21B group and negative control groups. The Celigo cell count and CCK-8 results showing the proliferation of the 143B **(B–D)** and U2-OS **(E–G)** cells in the shKIF21B group and negative control groups after treatment with IGF-1 (200 μg/L) for 24 h. Scale bar = 200 μm, *P < 0.05.

the PI3K/AKT pathway but also in the focal adhesion, MAPK, TNF, and ECM processes. We then examined the expression levels of PI3K/AKT-related proteins and apoptosis-related proteins. Knocking down KIF21B results in the downregulation of PI3K/AKT-related proteins and Bcl-2 and the upregulation of Bax. In addition, PI3K/AKT pathway agonists can reverse the regulatory effect of KIF21B on the proliferation and apoptosis of osteosarcoma cells.

Prior to this study, there was no clear evidence showing a relationship between KIF21B and PI3K/AKT. The PI3K/AKT signaling pathway has been shown to play a role in the regulation of kinesin protein (32, 33), but the detailed mechanism by which the PI3K/AKT pathway regulates kinesin is currently rarely studied. On the other hand, according to the results of the gene chip and bioinformatics, cell adhesion and ECM processes were also enriched in KIF21B-silenced cells. Since ECM

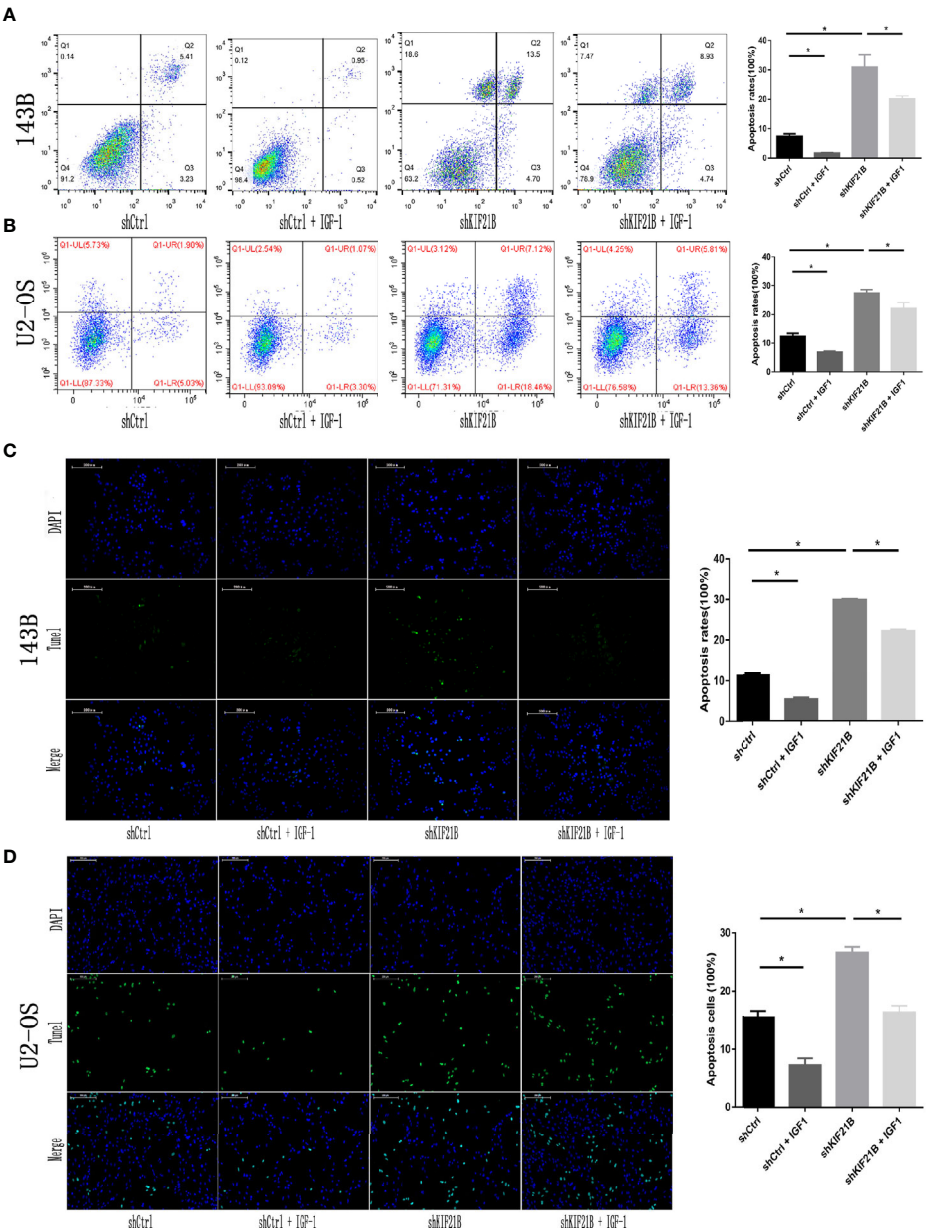


FIGURE 8 | Effects of IGF-1 on cell apoptosis after silencing KIF21B. Flow cytometry **(A, B)** and TUNEL assays **(C, D)** results showing the apoptosis rate in each group after treatment with IGF-1 (200 μg/L) for 24 h. Scale bar = 200 μm, *P < 0.05.

processes and cell adhesion can have a broad relationship with the PI3K/AKT pathway (34–37), we hypothesize that the mechanism by which KIF21B regulates osteosarcoma cells through the PI3K/AKT pathway is probably related to ECM processes and cell adhesion.

It should be noted that there are still several limitations in this study. First, we strictly selected participants to control for potential heterogeneity, which led to a small number of patients (17 patients). Second, the research on the cellular functions and molecular mechanisms of KIF21B was not comprehensive, and further studies are necessary to verify the role of KIF21B in the development and progression of OS. In summary, this experiment first confirmed that KIF21B, as an oncogene, plays an important role in the occurrence and development of OS. KIF21B silencing can also lead to the inhibition of cell proliferation and the induction of apoptosis. The mechanism by which KIF21B regulates OS might be associated with the persistent activation of the PI3K/AKT signaling pathway. However, the basic and clinical research investigating the role of KIF21B is still in the initial stages, and

the molecular mechanism involved in the occurrence and development of OS warrants further study.

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 studies involving human participants were reviewed and approved by the Ethics Committee of the Zhujiang Hospital of Southern Medical University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Ethics Committee of the Zhujiang Hospital of Southern Medical University. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR CONTRIBUTIONS

YD: Conceptualization, investigation, methodology, project administration, resources. SN and JL: Writing—review and editing, data curation, formal analysis, writing—original draft. SQ, KG, and YX: Formal analysis. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.606765/full#supplementary-material

SUPPLEMENTARY FIGURE 1 | Principal component analysis (PCA) of abnormal samples.

REFERENCES

1. Anninga JK, Gelderblom H, Fiocco M, Kroep JR, Taminiau AH, Hogendoorn PC, et al. Chemotherapeutic adjuvant treatment for osteosarcoma: where do we stand? *Eur J Cancer* (2011) 47:2431–45. doi: 10.1016/j.ejca.2011.05.030
2. Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. *Nat Rev Cancer* (2014) 14:722–35. doi: 10.1038/nrc3838
3. Lin AY, Kotova S, Yanagawa J, Elbuluk O, Wang G, Kar N, et al. Risk stratification of patients undergoing pulmonary metastasectomy for soft tissue and bone sarcomas. *J Thoracic Cardiovasc Surg* (2015) 149:85–92. doi: 10.1016/j.jtcvs.2014.09.039
4. Tirtei E, Asaftei SD, Manicone R, Cesari M, Paioli A, Rocca M, et al. Survival after second and subsequent recurrences in osteosarcoma: a retrospective multicenter analysis. *Tumori* (2018) 104:202–6. doi: 10.1177/0300891617753257
5. Marszalek JR, Weiner JA, Farlow SJ, Chun J, Goldstein LS. Novel dendritic kinesin sorting identified by different process targeting of two related kinesins: KIF21A and KIF21B. *J Cell Biol* (1999) 145:469–79. doi: 10.1083/jcb.145.3.469
6. Huang X, Liu F, Zhu C, Cai J, Wang H, Wang X, et al. Suppression of KIF3B expression inhibits human hepatocellular carcinoma proliferation. *Digest Dis Sci* (2014) 59:795–806. doi: 10.1007/s10620-013-2969-2
7. Moores CA, Milligan RA. Lucky 13-microtubule depolymerisation by kinesin-13 motors. *J Cell Sci* (2006) 119:3905–13. doi: 10.1242/jcs.03224
8. Walczak CE, Gayek S, Ohi R. Microtubule-depolymerizing kinesins. *Annu Rev Cell Dev Biol* (2013) 29:417–41. doi: 10.1146/annurev-cellbio-101512-122345
9. Zhao HQ, Dong BL, Zhang M, Dong XH, He Y, Chen SY, et al. Increased KIF21B expression is a potential prognostic biomarker in hepatocellular carcinoma. *World J Gastrointestinal Oncol* (2020) 12:276–88. doi: 10.4251/wjgo.v12.i3.276
10. Chen J, Yue C, Xu J, Zhan Y, Zhao H, Li Y, et al. Downregulation of receptor tyrosine kinase-like orphan receptor 1 in preeclampsia placenta inhibits human trophoblast cell proliferation, migration, and invasion by PI3K/AKT/mTOR pathway accommodation. *Placenta* (2019) 82:17–24. doi: 10.1016/j.placenta.2019.05.002
11. Josephs DH, Sarker D. Pharmacodynamic Biomarker Development for PI3K Pathway Therapeutics. *Trans Oncogenom* (2015) 7:33–49. doi: 10.4137/TOG.S30529
12. Polivka J Jr., Janku F. Molecular targets for cancer therapy in the PI3K/AKT/mTOR pathway. *Pharmacol Ther* (2014) 142:164–75. doi: 10.1016/j.pharmthera.2013.12.004
13. Faes S, Dormond O. PI3K and AKT: Unfaithful Partners in Cancer. *Int J Mol Sci* (2015) 16:21138–52. doi: 10.3390/ijms160921138
14. Noseda R, Guerrero-Valero M, Alberizzi V, Previtali SC, Sherman DL, Palmisano M, et al. Kif13b Regulates PNS and CNS Myelination through the Dlg1 Scaffold. *PloS Biol* (2016) 14:e1002440. doi: 10.1371/journal.pbio.1002440
15. Gao Y, Zheng H, Li L, Zhou C, Chen X, Zhou X, et al. KIF3C Promotes Proliferation, Migration, and Invasion of Glioma Cells by Activating the PI3K/AKT Pathway and Inducing EMT. *BioMed Res Int* (2020) 2020:6349312. doi: 10.1155/2020/6349312
16. Picci P. Osteosarcoma (osteogenic sarcoma). *Orphanet J Rare Dis* (2007) 2:6. doi: 10.1186/1750-1172-2-6
17. Jaffe N. Osteosarcoma: review of the past, impact on the future. The American experience. *Cancer Treat Res* (2009) 152:239–62. doi: 10.1007/978-1-4419-0284-9_12
18. Bertoni F, Bacchini P, Fabbri N, Mercuri M, Picci P, Ruggieri P, et al. Osteosarcoma. Low-grade intraosseous-type osteosarcoma, histologically resembling parosteal osteosarcoma, fibrous dysplasia, and desmoplastic fibroma. *Cancer* (1993) 71:338–45. doi: 10.1002/1097-0142(19930115)71:2<338::AID-CNCR2820710212>3.0.CO;2-H
19. Anderson CA, Massey DC, Barrett JC, Prescott NJ, Tremelling M, Fisher SA, et al. Investigation of Crohn's disease risk loci in ulcerative colitis further defines their molecular relationship. *Gastroenterology* (2009) 136:523–9 e3. doi: 10.1053/j.gastro.2008.10.032
20. Goris A, Boonen S, D'Hooghe MB, Dubois B. Replication of KIF21B as a susceptibility locus for multiple sclerosis. *J Med Genet* (2010) 47:775–6. doi: 10.1136/jmg.2009.075911
21. Yang X, Li M, Wang L, Hu Z, Zhang Y, Yang Q. Association of KIF21B genetic polymorphisms with ankylosing spondylitis in a Chinese Han population of Shandong Province. *Clin Rheumatol* (2015) 34:1729–36. doi: 10.1007/s10067-014-2761-5
22. Kreft KL, van Meurs M, Wierenga-Wolf AF, Melief MJ, van Strien ME, Hol EM, et al. Abundant kif21b is associated with accelerated progression in neurodegenerative diseases. *Acta Neuropathol Commun* (2014) 2:144. doi: 10.1186/s40478-014-0144-4
23. Olson HE, Shen Y, Poduri A, Gorman MP, Dies KA, Robbins M, et al. Microduplications of 1q32.1 associated with neurodevelopmental delay. *Eur J Med Genet* (2012) 55:145–50. doi: 10.1016/j.ejmg.2011.12.008
24. Liu X, Gong H, Huang K. Oncogenic role of kinesin proteins and targeting kinesin therapy. *Cancer Sci* (2013) 104:651–6. doi: 10.1111/cas.12138
25. Yokota K, Sasaki H, Okuda K, Shimizu S, Shitara M, Hikosaka Y, et al. KIF5B/RET fusion gene in surgically-treated adenocarcinoma of the lung. *Oncol Rep* (2012) 28:1187–92. doi: 10.3892/or.2012.1908

26. Noguchi S, Kumazaki M, Yasui Y, Mori T, Yamada N, Akao Y. MicroRNA-203 regulates melanosome transport and tyrosinase expression in melanoma cells by targeting kinesin superfamily protein 5b. *J Invest Dermatol* (2014) 134:461–9. doi: 10.1038/jid.2013.310
27. O'Hare M, Shadmand M, Sulaiman RS, Sishtla K, Sakisaka T, Corson TW. Kif14 overexpression accelerates murine retinoblastoma development. *Int J Cancer* (2016) 139:1752–8. doi: 10.1002/ijc.30221
28. Corson TW, Zhu CQ, Lau SK, Shepherd FA, Tsao MS, Gallie BL. KIF14 messenger RNA expression is independently prognostic for outcome in lung cancer. *Clin Cancer Res An Off J Am Assoc Cancer Res* (2007) 13:3229–34. doi: 10.1158/1078-0432.CCR-07-0393
29. Corson TW, Gallie BL. KIF14 mRNA expression is a predictor of grade and outcome in breast cancer. *Int J Cancer* (2006) 119:1088–94. doi: 10.1002/ijc.21954
30. Madhavan J, Mitra M, Mallikarjuna K, Pranav O, Srinivasan R, Nagpal A, et al. KIF14 and E2F3 mRNA expression in human retinoblastoma and its phenotype association. *Mol Vision* (2009) 15:235–40.
31. Wordeman L. How kinesin motor proteins drive mitotic spindle function: Lessons from molecular assays. *Semin Cell Dev Biol* (2010) 21:260–8. doi: 10.1016/j.semcdb.2010.01.018
32. Liu Z, Rebowe RE, Wang Z, Li Y, Wang Z, DePaolo JS, et al. KIF3a promotes proliferation and invasion via Wnt signaling in advanced prostate cancer. *Mol Cancer Res MCR* (2014) 12:491–503. doi: 10.1158/1541-7786.MCR-13-0418
33. Tanaka Y, Niwa S, Dong M, Farkhondeh A, Wang L, Zhou R, et al. The Molecular Motor KIF1A Transports the TrkA Neurotrophin Receptor and Is Essential for Sensory Neuron Survival and Function. *Neuron* (2016) 90:1215–29. doi: 10.1016/j.neuron.2016.05.002
34. Zhu X, Efimova N, Arnette C, Hanks SK, Kaverina I. Podosome dynamics and location in vascular smooth muscle cells require CLASP-dependent microtubule bending. *Cytoskeleton* (2016) 73:300–15. doi: 10.1002/cm.21302
35. Sun L, Dong Y, Zhao J, Yin Y, Zheng Y. The CLC-2 Chloride Channel Modulates ECM Synthesis, Differentiation, and Migration of Human Conjunctival Fibroblasts via the PI3K/Akt Signaling Pathway. *Int J Mol Sci* (2016) 17:910. doi: 10.3390/ijms17060910
36. Qu H, Sun H, Wang X. Neogenin-1 Promotes Cell Proliferation, Motility, and Adhesion by Up-Regulation of Zinc Finger E-Box Binding Homeobox 1 Via Activating the Rac1/PI3K/AKT Pathway in Gastric Cancer Cells. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol* (2018) 48:1457–67. doi: 10.1159/000492255
37. Luo J, Yao JF, Deng XF, Zheng XD, Jia M, Wang YQ, et al. 15-EET induces breast cancer cell EMT and cisplatin resistance by up-regulating integrin alphavbeta3 and activating FAK/PI3K/AKT signaling. *J Exp Clin Cancer Res CR* (2018) 37:1423. doi: 10.1186/s13046-018-0694-6

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 Ni, Li, Qiu, Xie, Gong and Duan. 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 Oncology

ORIGINAL RESEARCH
published: 29 January 2021
doi: 10.3389/fonc.2020.566804

Proteomic Alterations in Multiple Myeloma: A Comprehensive Study Using Bone Marrow Interstitial Fluid and Serum Samples

*Venkatesh Chanukuppa[1,2], Ravindra Taware[1], Khushman Taunk[1], Tathagat Chatterjee[3], Sanjeevan Sharma[4], Venkatesan Somasundaram[4], Faraz Rashid[5], Dipankar Malakar[5], Manas K. Santra[6] and Srikanth Rapole[1]**

[1] Proteomics Lab, National Centre for Cell Science, Pune, India, [2] Savitribai Phule Pune University, Pune, India, [3] Army Hospital (Research and Referral), New Delhi, India, [4] Armed Forces Medical College, Pune, India, [5] Sciex, Gurgaon, India, [6] Cancer Biology and Epigenetics Lab, National Centre for Cell Science, Pune, India

OPEN ACCESS

Edited by:
Suman S. Thakur,
Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Hong Jin,
Fudan University, China
Amol Ratnakar Suryawanshi,
Institute of Life Sciences (ILS), India

***Correspondence:**
Srikanth Rapole
rsrikanth@nccs.res.in

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 28 May 2020
Accepted: 07 December 2020
Published: 29 January 2021

Citation:
Chanukuppa V, Taware R, Taunk K, Chatterjee T, Sharma S, Somasundaram V, Rashid F, Malakar D, Santra MK and Rapole S (2021) Proteomic Alterations in Multiple Myeloma: A Comprehensive Study Using Bone Marrow Interstitial Fluid and Serum Samples. Front. Oncol. 10:566804. doi: 10.3389/fonc.2020.566804

Multiple myeloma (MM) is a plasma cell-associated cancer and exists as the second most common hematological malignancy worldwide. Although researchers have been working on MM, a comprehensive quantitative Bone Marrow Interstitial Fluid (BMIF) and serum proteomic analysis from the same patients' samples is not yet reported. The present study involves the investigation of alterations in the BMIF and serum proteome of MM patients compared to controls using multipronged quantitative proteomic approaches *viz*., 2D-DIGE, iTRAQ, and SWATH-MS. A total of 279 non-redundant statistically significant differentially abundant proteins were identified by the combination of three proteomic approaches in MM BMIF, while in the case of serum 116 such differentially abundant proteins were identified. The biological context of these dysregulated proteins was deciphered using various bioinformatic tools. Verification experiments were performed in a fresh independent cohort of samples using immunoblotting and mass spectrometry based SRM assays. Thorough data evaluation led to the identification of a panel of five proteins *viz*., haptoglobin, kininogen 1, transferrin, and apolipoprotein A1 along with albumin that was validated using ELISA in a larger cohort of serum samples. This panel of proteins could serve as a useful tool in the diagnosis and understanding of the pathophysiology of MM in the future.

Keywords: multiple myeloma, multipronged quantitative proteomics, bone marrow interstitial fluid, serum biomarkers, SWATH-MS, 2D-DIGE, iTRAQ

INTRODUCTION

Multiple myeloma (MM) accounts for 13% of all hematological malignancies across the world and remains an incurable disease with the worst prognosis (1). The incidence rate of MM varies greatly throughout the world and in developing countries like India, four out of 100,000 individuals are diagnosed with this deadly malignant disease every year (2). MM is characterized by clonal expansion of plasma cells and their subsequent accumulation in the bone marrow which further

leads to bone resorption owing to overproduction of antibodies (3). The monoclonal gammopathy of undetermined clinical significance (MGUS) is a benign form of MM that could lead to smoldering myeloma and finally progress into symptomatic myeloma (4). MM is a heterogeneous disease and has several factors attributed to its pathology like hyperdiploidy of specific chromosomes such as 3, 5, 7, 9, 11, 15, 19 and translocations of t (11; 14) (q13; q32) (5). The current clinical practices largely rely on genetic, immunohistochemical and flow cytometric analysis as well as imaging techniques like MRI and PET/CT scan for diagnosis and the evaluation of progression of MM (6). Moreover, researchers also proposed an international staging system for MM using the serum albumin and Beta-2-microglobulin (B2M) levels (7). In addition to these two protein markers, the staging system was further refined with fluorescent *in situ* hybridization technique as well (7, 8).

Interestingly, there is no clinically relevant protein marker yet available for MM diagnosis (6). The current diagnostic methods such as imaging, cytogenetic analysis or immunohistochemistry are expensive, time-consuming, tedious and require highly skilled manpower along with sophisticated equipment. Therefore, there is an unmet need for the biomarkers which can be employed in a clinical setting with high confidence and protein-based biomarkers holds tremendous potential towards it. Moreover, a combination of the routinely used markers such as B2M and serum albumin with newly identified biomarkers panel could be helpful for enhanced specificity and accuracy of MM diagnosis. Mass spectrometry (MS) based proteomics serves as an excellent approach that provides information about protein alterations and modifications in a variety of human samples including the patient's tissues and body fluids (9). Globally, many research groups have explored and identified potential biomarkers and targets for various cancers using MS-based proteomic approaches (10–14). However, in the case of MM, limited studies have explored the identification of potential candidate markers using proteomic approaches (15–21). Zhang et al. analyzed the pooled serum samples of MM patients towards the identification of markers using liquid chromatography tandem MS/MS (22). In another study, Wang et al. built MM serum diagnostic model and pattern recognition software by using the magnetic bead-based MALDI-TOF MS (23). Lu et al. identified high abundant proteins in MM cells through 2-DE followed by MALDI-TOF MS/MS analysis and the majority of these proteins were categorized under cytoskeletal, chaperone, oxidoreductase and protease class (24). Likewise, Ma and co-authors identified 11 differentially expressed proteins in MM serum using 2-DE followed by MALDI-TOF MS/MS analysis (21). Recently, Bai et al. showed the correlation of fibrinogen alpha chain, dihydropyrimidinase-like 2, platelet factor 4 and alpha-fetoprotein proteins with disease states of MM (20). Apart from that, other proteomic studies reported on MM, so far focused on the identification of proteomic alterations in response to drug treatments (25–29).

To identify the suitable biomarker for MM, the choice of biospecimen for proteomic investigation is very important. Being a proximal biofluid for the hematological malignancies, bone marrow interstitial fluid (BMIF) could serve as a potential source for identifying diagnostic, prognostic and therapeutic markers for various blood cancers including MM (30). On the other hand, blood serum serves as a minimally invasive biofluid and is routinely used in clinical diagnosis. The majority of reported and clinically approved diagnostic markers are serum proteins. Therefore, in this study, we have investigated the proteome alterations in the MM BMIF and serum through multipronged quantitative proteomics by employing two-dimensional difference gel electrophoresis (2D-DIGE), isobaric tags for relative and absolute quantitation (iTRAQ) and Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS) approaches. The statistically significant proteins from both the biological matrices were also compared with each other to identify the common proteins with a similar trend of differential abundance. Additionally, the probable biological and functional roles of the differentially abundant proteins were probed in the manifestation of MM disease pathology by using various *in silico* tools. Furthermore, a selected panel of statistically significant proteins was verified for their differential abundance by immunoblotting as well as MS-based SRM assays. Moreover, potential candidate proteins were also validated in a fresh independent cohort of serum samples using Enzyme-linked immunosorbent assay (ELISA), as it is a minimal invasive fluid that could easily be used for the diagnostics applications. The significance of this study remains in the fact that the proteins which are observed in the BMIF and also reflected in the serum with similar expression profile are proposed as a potential candidate biomarker panel for MM diagnosis. Though BMIF serves as proximal fluid and an excellent source of information on disease pathophysiology, it is an extremely invasive sampling procedure, which puts a lot of stress on the suspected patients of MM. Therefore, the panel of protein markers which are present both in BMIF and serum with a similar expression profile could serve as an excellent alternative and complementary markers with high specificity and accuracy. To the best of our knowledge, this is the first comprehensive proteomics study involving MM clinical bone marrow interstitial fluid and serum samples which utilizes multipronged proteomic approaches.

MATERIALS AND METHODS

Sample Collection

The Institutional Ethics Committee of Armed Forces Medical College, Pune and National Centre for Cell Science, Pune approved this study. A total of 156 volunteers (MM = 64, non-hematological malignancies control = 28, healthy controls = 64) were recruited for the study. Due to the unavailability of bone marrow interstitial fluid from healthy controls, we included the non-hematological malignancy samples in this study as BMIF control samples as these samples are devoid of any hematological malignancies. All the volunteers were informed about the study and written informed consent was obtained from each of them before clinical sample collection. The bone marrow aspirate and

peripheral blood samples were simultaneously collected from the same study cohort. Freshly diagnosed naive MM patients without any other malignancy, diabetes and hypertension were recruited for this study. Similarly, controls without any ailments such as diabetes or hypertension were recruited for this study. The average age of the MM subjects recruited for this study was 63.95 years and that of healthy volunteers and non-hematological malignancy patients was 57.54 and 56.2 years, respectively (**Table S1**). Parameters such as serum creatinine and M protein concentrations were considered for discriminating the early stage of MM samples. Total 24 BMIF and 24 serum samples (MM = 12, control = 12) were utilized for the discovery phase. For 2D-DIGE and iTRAQ, 3 pooled samples were used for each set and 4 sets of experiments were performed. Whereas 12 individual samples were used for SWATH-MS experiments and each sample was acquired in triplicate. A separate cohort of samples of MM patients (n = 16) and respective controls (n = 16) was used for WB and SRM-based verification assays. The verification experiments included a different set of 8 controls and 8 MM samples for all the western blotting experiments. Similarly, the SRM experiments comprised of 8 biological replicates from the different verification sample set. Finally, in the validation phase, we performed ELISA experiments on serum samples obtained from a separate cohort of 72 volunteers (MM = 36, control = 36). A flow chart depicting the design of the experiment and sample allotment strategy is shown in **Figure 1**.

Protein Extraction and Sample Preparation

The bone marrow aspirate and blood samples were incubated for 30 min at room temperature and then centrifuged at 3,000 rpm for 10 min at 10°C to separate the BMIF and serum respectively. The BMIF and serum samples were transferred in sterile cryovials, labeled and stored at −80°C freezer until further experiments. BMIF and serum samples were depleted for most abundant proteins such as albumin and IgG by using affinity chromatography based spin trap format depletion kit and desalted using a 2D clean-up kit (both GE Healthcare, USA) as per instructions provided by the manufacturer. BMIF and serum protein pellets, thus obtained, were dissolved in rehydration buffer (7 M Urea, 2 M Thiourea and 2% CHAPS) and the protein estimation was performed by a 2D Quant kit (GE Healthcare, USA).

Cy Dye Labeling and 2D-DIGE Analysis

The 2D-DIGE experiments were performed as per the protocol reported previously by Rapole and co-authors (31). Briefly, 60 µg of BMIF and serum proteins from controls and MM patients were labeled with 400 pmol of CyDyes Cy5, Cy3 respectively. Moreover, an internal control having 1:1 concentration of both samples was constituted and labeled with Cy2. The labeled samples were rehydrated on a 24 cm linear IPG strip (pH 4–7), and the strips were subjected to isoelectric focusing (IEF) through EttanIPGphor 3 instrument (GE Healthcare, USA).

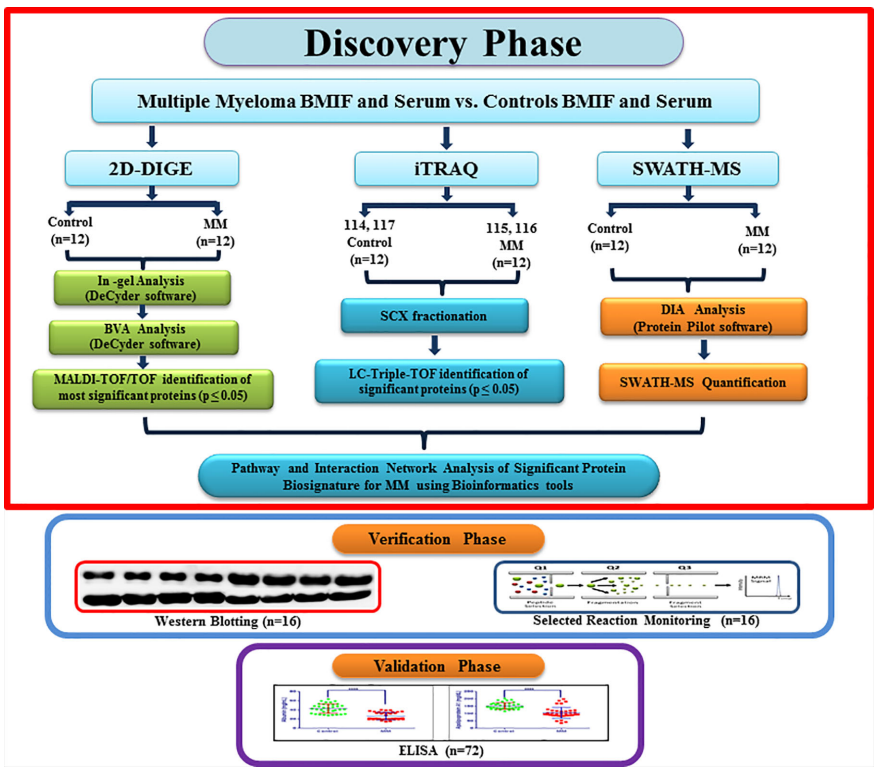


FIGURE 1 | A flowchart depicting experimental design along with the sample allocation strategy for Multiple myeloma BMIF and serum proteomics.

Following the first dimension IEF procedure, the second-dimension SDS-PAGE was performed and the gels were scanned through the Typhoon FLA 9500 biomolecular imager (GE Healthcare, USA). The images were analyzed using DeCyder 2D software; version 7.0 (GE Healthcare, USA) for both the differential in-gel analysis and biological variation analysis (BVA). The statistically significant protein spots passing the t-test (p-value ≤ 0.05) and present in all the gels were selected for MALDI-TOF/TOF identification.

In-Gel Digestion and Protein Identification Using MALDI-TOF/TOF

Differentially expressed protein spots identified in 2D-DIGE were excised automatically through Ettan spot picker (GE Healthcare, USA) from the preparative gel and subjected to in-gel digestion of the proteins as reported earlier (32). Briefly, the gel pieces were reduced with DTT and alkylated with IAA and subjected to enzymatic digestion using trypsin (100 ng) by keeping at overnight incubation at 37°C. The tryptic peptides from in-gel digestion were extracted and spotted on a MALDI plate with CHCA matrix (10 mg/ml). Protein identification was performed using a 4,800 MALDI-TOF/TOF mass spectrometer (AB Sciex, USA) linked to a 4,000 series explorer software (v.3.5.3) equipped with Nd: YAG 355 nm laser and a repetition rate of 200. The mass range of 800 to 4,000 Da was used, and the mass spectra were acquired in reflector mode with 20 kV and 18 kV as acceleration and extraction voltages, respectively. MS/MS spectra were acquired for the 12 most abundant precursor ions, with a total accumulation of 2,500 laser shots and collision energy of 2 kV. The MASCOT version 2.1 (http://www.matrixscience.com) was used for the data analysis by keeping the taxonomy as *Homo sapiens*, database as SwissProt, enzyme as trypsin, oxidation of methionine as variable and carbamidomethylation of cysteine as a fixed modification. The MS mass tolerance was set as 75 ppm and MS/MS as 0.4 Da.

iTRAQ Labeling and Prefractionation

Hundred microgram protein from MM BMIF and serum [three samples pooled (one set), total four sets] and respective controls [three samples pooled (one set), total four sets] were subjected to in-solution digestion using trypsin. The four plex iTRAQ labeling of digested peptides was performed as per the manufacturer's protocol (SCIEX, USA). Briefly, iTRAQ reagents dissolved in ethanol were added to the respective protein sample i.e., 114-controls, 115-MM samples, 116-MM samples and 117-controls and incubated at RT for 1 h. The labeled samples were pooled and concentrated using SpeedVac (Savant- SPD 1010, Thermo Electron Corporation, USA). Labeled peptide sample was fractionated by SCX chromatography (Poly-SULFOETHYL A column, 100 × 4.6 mm, 5 mm, 300 Å, PolyLC, Columbia, MD) using a Shimadzu HPLC. The fractionated peptide samples were again concentrated under vacuum and desalted using C18 ZipTips (Millipore, USA) before performing the LC-MS/MS analysis. The desalted peptide samples were separated and analyzed through Eksigent MicroLC 200 System (Eksigent, Canada) which was coupled to a Triple TOF 5600 high-resolution mass spectrometer (SCIEX, USA). The peptides were separated using a linear gradient of 7–35% ACN for 60 min for peptide elution from the analytical column with a flow rate of 8 µl/min. The column temperature was set as 40°C, and the autosampler temperature was kept as 4°C. The mass spectrometric analysis was executed with m/z ranging from 100–3,200, MS scan rate of six spectra, MS/MS scan rate of three spectra, with top 15 peaks for MS/MS analysis. The protein identification and quantitation were performed by ProteinPilot software (version 4.0, SCIEX) through the SwissProt database with 1% FDR and one missed cleavage as input parameters.

SWATH-MS Analysis

The label-free SWATH-MS analysis was performed on BMIF and serum samples (12 MM and 12 controls), which were depleted for Albumin and IgG. An equal amount of BMIF and serum proteins were subjected to trypsin digestion, and the peptides were analyzed using a Triple TOF 5600 mass spectrometer (SCIEX, USA) equipped with Eksigent Nano 2D Ultra 2D plus (Eksigent, Canada) having an Eksigent Nano LC 3C18 CL reverse phase column (75 µm × 15 mm, 3 µm, 120 Å) along with NanoLC Trap Chrom XP C-18-CL (3 µm, 120 Å, 350 µm × 0.5 mm) column. Data dependent analysis (DDA) was performed for the individual samples to generate high quality spectral ion libraries for SWATH-MS analysis, by operating the mass spectrometer with specific parameters as mentioned elsewhere (33). Technical triplicates of SWATH-MS were carried out for each BMIF and serum samples. For SWATH-MS experiments, the instrument was tuned to optimize the quadrupole settings for the selection of the precursor ion window of 25 m/z width. An isolation width of 26 m/z with 1 m/z for the window overlap was used and a set of 34 overlapping windows was built for covering the precursor mass range of 400–1,250 m/z. The SWATH-MS/MS spectra were derived from 100 m/z to 2,000 m/z. The ions were fragmented in the collision cell using rolling collision energy (CE) with an additional CE spread of ±15 eV. A dwell time of 96 ms was used for all fragment-ion scans in high-sensitivity mode, and for each SWATH-MS cycle, a survey scan in high-resolution mode was acquired for 100 ms, resulting in a duty cycle of 3.33 s. The dual-source parameters were as follows: ion source gases GS1, GS2, curtain gas at 25 psi, temperature 200°C, and ion spray voltage floating at 5,500 V. High-quality spectral ion libraries were generated for SWATH analysis through DDA of individual samples. The peak extraction and spectral alignment were performed using the Peakview software (version 2.2, SCIEX, USA) with the following specific parameters: no. of peptides two, no. of transitions 10, peptide confidence 99%, XIC width 30 ppm, XIC extraction window 3 min. The data was further subjected to MarkerView software (version 1.3.1, SCIEX, USA) to get statistical data interpretation. In MarkerView, the results were shown as three output files containing AUC of the ions, the summed intensity of peptides for protein and the summed intensity of ions for the peptide. The summed intensity of peptides was used for the further relative quantitative and

multivariate statistical analysis using MetaboAnalyst 3.0 and SIMCA 14.1 platforms. The statistical significance was tested using the student's t-test and FDR corrections. The reproducibility of the data was checked using retention time matching of some of the selected peptide spectra. The data obtained from the proteomic analysis were subjected to mathematical normalization in order to obtain the Gaussian distribution. The data were normalized to constant sum, cube root transformed and auto-scaled using MetaboAnalyst web application.

Functional and Interaction Analysis Using Bioinformatics Tools

Proteins that were significantly differentially expressed (p-value < 0.05 and FC ≥ 1.5/≤ 0.67) were interpreted using Database for Annotation, Visualization and Integrated Discovery (DAVID version 6.8, http://david.abcc.ncifcrf.gov/home.jsp), Protein Analysis Through Evolutionary Relationships (PANTHER version 11.1, http://www.pantherdb.org/) analysis and Ingenuity Pathway Analysis (IPA). DAVID analysis was performed by selecting functional annotation, Uniprot accession as identifier ID and extracted the biological functions. Similarly, PANTHER analysis was performed by keeping the *Homo sapiens* as a selected organism. Functional classification viewed in graphic charts for select analysis and exported the molecular function, biological process, cellular component, protein class and pathways. Likewise, *in silico* bioinformatics analysis was performed by Ingenuity Pathway Analysis (IPA) software (Qiagen Bioinformatics, India) to identify the Canonical pathways, upstream regulators and toxicology functions.

Western Blot Based Verification

Western blot analyses of BMIF and serum samples (MM = 8, Control = 8) were performed as described previously (34). Briefly, BMIF and serum proteins were separated on a 12% SDS PAGE gel (40 µg per well) and then transferred onto a nitrocellulose membrane under semi-dry conditions using ECL semi-dry transfer unit (GE Healthcare, USA). Western blot was performed with a monoclonal/polyclonal antibody against Haptoglobin, Kininogen1, Alpha 1 antitrypsin, Gelsolin, Zinc-2-alpha glycoprotein, Apolipoprotein A1, Transferrin, β-actin (all from Sigma Aldrich, USA). An appropriate secondary antibody conjugated with horseradish peroxidase enzyme (HRP) (GE Healthcare, USA) was employed to detect the respective primary antibody. After treating with a chemiluminescent substrate, the blots were visualized by ImageQuant LAS 4000 instrument (GE Healthcare, USA).

Selected Reaction Monitoring Based Verification

The differentially expressed BMIF and serum proteins were further cross-verified by SRM assays using a new cohort of samples (MM = 8, control = 8). Transitions for each of the proteins to be verified through the SRM experiments were established *via in silico* approach on SKYLINE 3.1 software, cross-referred from SRM Atlas, MRM based public libraries and literature (33, 35, 36). Tryptic peptides with an average length of 5–10 amino acids, devoid of residues prone to post-translational modifications were chosen for establishing SRM transitions. Unique peptides for each protein were taken as a target and the collision energy obtained from SKYLINE software was used. The unique peptide transition list was imported to 4000 Q-TRAP LC-MS/MS system (SCIEX, USA) attached to an Eksigent Micro LC System. SB C18 micro LC column (0.3 mm × 100 mm, 300 Å pore size, 5 µm particle size) from Agilent was used for the separation of peptides with 6 µl flow rate. The raw data files were analyzed using Skyline software for further quantitation. All the spectra were manually interpreted for their differential expression.

Enzyme-Linked Immunosorbent Assay Based Validation

In the validation phase, we used a separate cohort of 72 volunteers (MM = 36, control = 36) to perform ELISA based validation in serum samples. All the ELISA experiments were performed according to the manufacturer's instructions. We validated a panel of five proteins with the Human Haptoglobin Quantikine ELISA kit (catalog #DHAPG0; R&D Systems, Inc., Minneapolis, USA), Human Kininogen DuoSet ELISA kit (catalog #DY1569-05; R&D Systems, Inc., Minneapolis, USA), Human Apolipoprotein A-I Quantikine ELISA Kit (catalog #DAPA10; R&D Systems, Inc., Minneapolis, USA), Human sTfR Quantikine IVD ELISA Kit (1 Kit) (catalog #DTFR1; R&D Systems, Inc., Minneapolis, USA), Human Serum Albumin DuoSet ELISA, 15 Plate (1 KIT) (catalog #DY1455; R&D Systems, Inc., Minneapolis, USA). The optical density of the each sample was determined by using an EPOCH 2 microplate reader (Biotek, USA) set to 450 nm and 570 nm. Optical imperfections correction was done by subtracting the readings of 570 nm from 450 nm readings. Further quantitative analysis was performed using GraphPad Prism software.

RESULTS

Proteome Alterations in MM BMIF and Serum Using Multipronged Proteomic Approaches

To identify the MM induced differentially expressed proteins in MM BMIF and serum samples, the multipronged quantitative proteomic approaches were employed to maximize the proteomic coverage. The proteins identified in both the biomatrices with similar differential regulation and statistical significance were projected as potential candidate biomarkers. The results obtained from the experiments are discussed henceforth.

Identification of BMIF and Serum Proteome Alterations in MM Using 2D-DIGE

Decyder software revealed 1,450–1,600 protein spots which emerged as differentially expressed proteins in MM BMIF as compared to control. Biological variations of the gels were compared against a master gel through the BVA module to obtain statistically significant protein spots. Finally, 186 protein

spots showed the statistical significance of p-value ≤0.05 (student's t-test) with ≥1.5 fold change for up-regulation and ≤0.67 as fold change for down-regulation. Out of these differentially expressed protein spots, 96 protein spots were up-regulated and 90 protein spots showed down-regulation as compared to control. A total of 112 protein spots were identified through the 4800 MALDI-TOF/TOF MS (Sciex), and 31 proteins were found to be non-redundant. Of these 31 non-redundant proteins, 19 proteins were found to be up-regulated, while 12 proteins showed decreased expression in MM as compared to control. The list of differentially expressed BMIF proteins identified using 2D-DIGE is mentioned in **Table S2**. A representative combined 2D-DIGE image with BVA views showing 3D differential expression of the selected protein spots is depicted in **Figure 2**.

Similarly, serum 2D-DIGE analysis revealed approximately 1,200–1,300 protein spots out of which 153 protein spots showed differential regulation in MM with criteria of statistical significance of p-value ≤0.05 and fold change ≥1.5/≤0.67. Out of these differentially expressed protein spots, 90 protein spots were up-regulated and 63 protein spots showed down-regulation. A total of 96 proteins were identified through the MALDI-TOF/TOF MS, and 29 proteins were found to be non-redundant. Of these non-redundant proteins, 18 proteins were up-regulated and 11 proteins were down-regulated in MM as compared to control. The list of differentially expressed serum proteins identified using 2D-DIGE is mentioned in **Table S3**. A representative combined 2D-DIGE image with BVA views showing 3D differential expression of the selected protein spots is depicted in **Figure 3**.

Identification of BMIF and Serum Proteome Alterations in MM Using iTRAQ

A total of 719 proteins were identified in BMIF using iTRAQ and out of these 183 proteins were found as differentially expressed with criteria of fold change ≥1.5/≤0.67. A total of 91 proteins showed up-regulation and 92 proteins showed down-regulation in MM BMIF as compared to the control samples. The complete list of significant differentially regulated BMIF proteins identified using iTRAQ methodology is provided in **Table S4**.

In serum, a total of 650 proteins were identified using iTRAQ and out of these, 68 proteins were found as differentially expressed based on the fold change criteria of ≥1.5/≤0.67. Among differentially expressed proteins, 30 proteins were up-regulated and 38 showed down-regulation in MM patients as compared to healthy controls. The complete list of significant differentially regulated serum proteins identified using iTRAQ methodology is provided in **Table S5**.

Identification of BMIF and Serum Proteome Alterations in MM Using SWATH-MS

SWATH-MS library was built for 1236 BMIF proteins and out of these 283 proteins were commonly identified in all samples. The statistical analysis revealed 104 proteins differentially expressed out of which 41 proteins showed a pattern of up-regulation and 63 proteins showed down-regulation in MM samples as compared to controls (**Table S6**). The differential expression threshold criteria were the same as mentioned in section *Identification of BMIF and Serum Proteome Alterations in MM Using 2D-DIGE* and *Identification of BMIF and Serum Proteome*

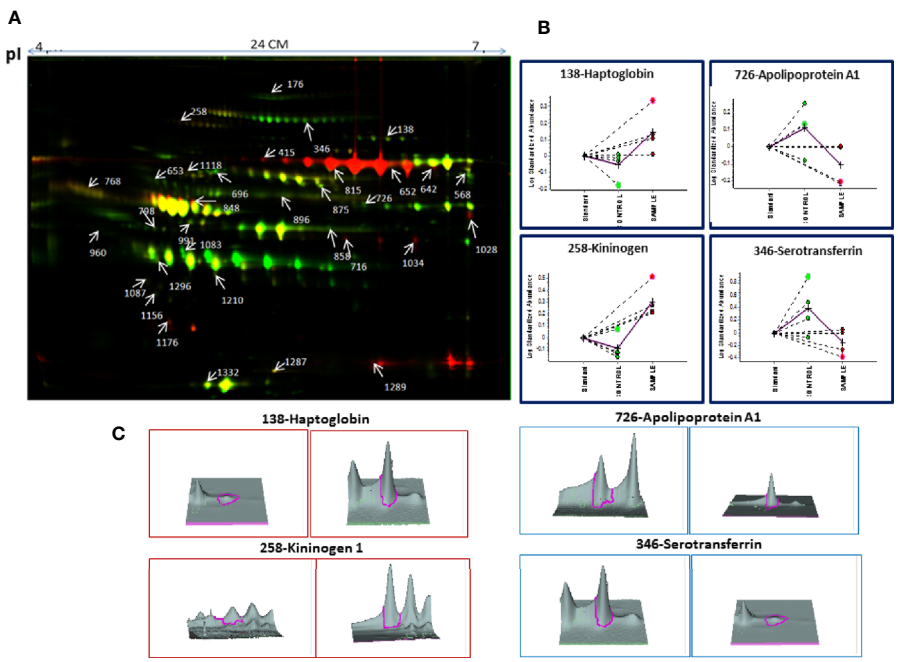


FIGURE 2 | **(A)** The representative 2D-DIGE image for MM BMIF *vs.* Control BMIF. **(B)** Dysregulation profile for some of the identified proteins obtained through BVA module of DeCyder software and **(C)** 3-D View of these proteins obtained through BVA module of DeCyder software.

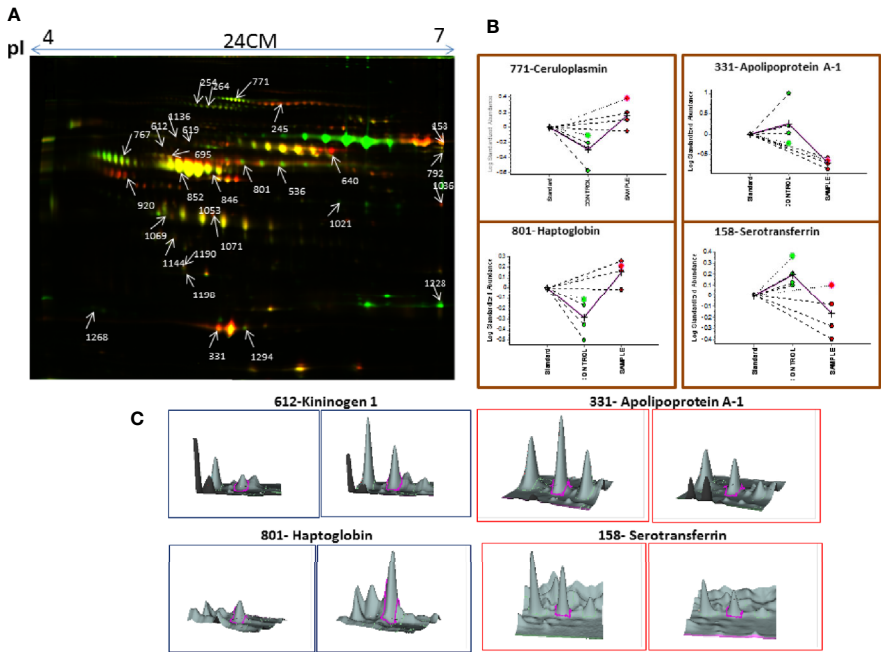


FIGURE 3 | **(A)** The representative 2D-DIGE image for MM serum *vs.* Control serum. **(B)** Dysregulation profile for some of the identified proteins obtained through BVA module of DeCyder software and **(C)** 3-D View of these proteins obtained through BVA module of DeCyder software.

Alterations in MM Using iTRAQ. Multivariate statistical analysis such as OPLS-DA for BMIF SWATH-MS data revealed good discrimination between the control and the MM groups (**Figure 4**).

Likewise, the SWATH-MS library was also built for a total of 1,184 serum proteins and out of these, 213 proteins were consistently observed in all samples. A total of 42 differentially expressed proteins were identified and out of which 18 proteins were up-regulated and 24 proteins showed down-regulation in MM samples as compared to controls (**Table S7**). The differential expression threshold criteria applied were the same as mentioned above. Furthermore, multivariate statistical analysis such as OPLS-DA revealed distinct clustering of MM and healthy control serum samples in score plots (**Figure 5**).

Identification of BMIF and Serum Proteome Alterations Using All Three Approaches

Since each of the quantitative proteomic approaches has its advantages and limitations, combining all these approaches in the discovery phase results in a maximum number of differentially expressed proteins. Using multipronged proteomic approaches, a total of 279 non-redundant BMIF proteins were found to be differentially regulated in MM (**Table S8**). Overall, among all of these differentially expressed proteins, 11 proteins were identified in both 2D-DIGE and iTRAQ, 17 proteins were observed in both iTRAQ and SWATH-MS, 14 proteins were common in both SWATH-MS and 2D-DIGE, and five proteins were presented in all the three approaches (**Figure 6A**). Interestingly, 11 proteins were observed only in DIGE, 159 proteins were specifically observed using the only iTRAQ, and 77 proteins were observed using only the SWATH-MS approach. The partial list of statistically significant differentially regulated proteins is summarized in **Table 1**.

Similarly, by using multipronged proteomic approaches, 116 non-redundant serum proteins were found to be differentially regulated in MM (**Table S9**). Out of the 116 proteins, 10 proteins were common in both 2D-DIGE and iTRAQ, 12 proteins were detected in both iTRAQ and SWATH-MS, five proteins were identified in both SWATH-MS and 2D-DIGE, and four proteins were observed in all the three approaches (**Figure 7A**). Eighteen proteins were observed only in DIGE, 50 proteins were specifically detected using only iTRAQ, and 29 proteins were identified using only SWATH-MS approach. The partial list of statistically significant and differentially regulated proteins is illustrated in **Table 2**. Identification of proteins in two or more approaches provides its definitive association with a particular disease, in this case MM.

Common Proteins Identified and Validated in MM BMIF and Serum Samples

Multipronged quantitative proteomic approaches yielded a total of 279 and 116 non-redundant differentially abundant proteins in BMIF and serum samples of MM study cohorts, respectively. A total of 55 proteins with statistically significant differential abundance were identified to be common in both MM BMIF and serum study cohorts. Interestingly, 41 proteins showed a similar pattern of differential expression and among these, 23 proteins were up-regulated and 18 proteins were down-regulated.

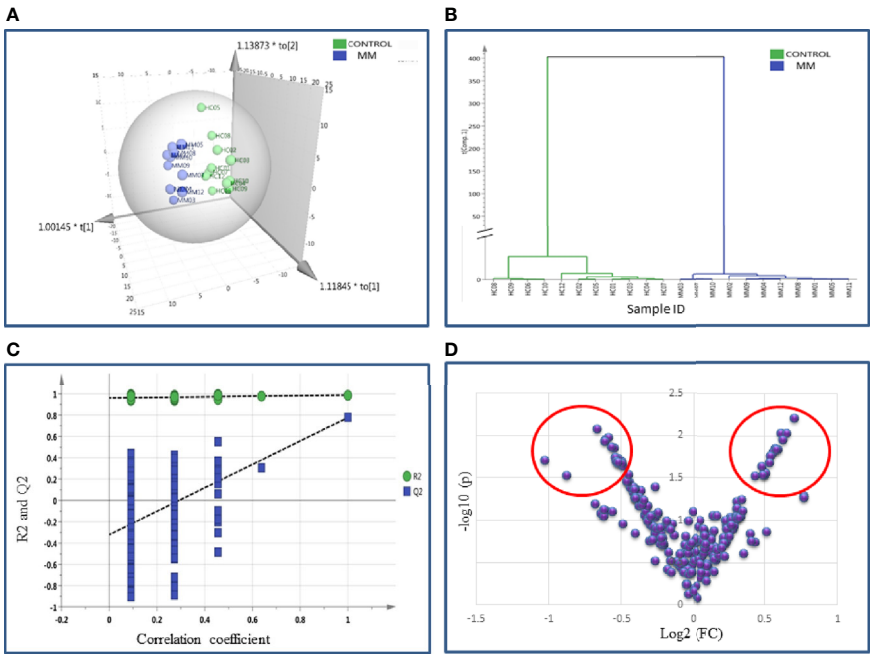


FIGURE 4 | Multivariate statistical analysis of MM BMIF *vs.* Control dataset using SWATH MS data. **(A)** OPLS-DA score plot for MM *vs.* Controls. **(B)** Dendrogram showing the hierarchical clustering between MM and controls based on DEPs. **(C)** Permutation test statistics of the OPLS-DA model of MM *vs.* Controls with Y-axis intercepts: R2 = (0.0, 0.894), Q2 = (0.0, −0.426), **(D)** V-Plot showing the most significant proteins in red circles. One sample in MM was removed from the SIMCA analysis as they had high degree of variation hence the data shown is for eleven MM samples and twelve controls.

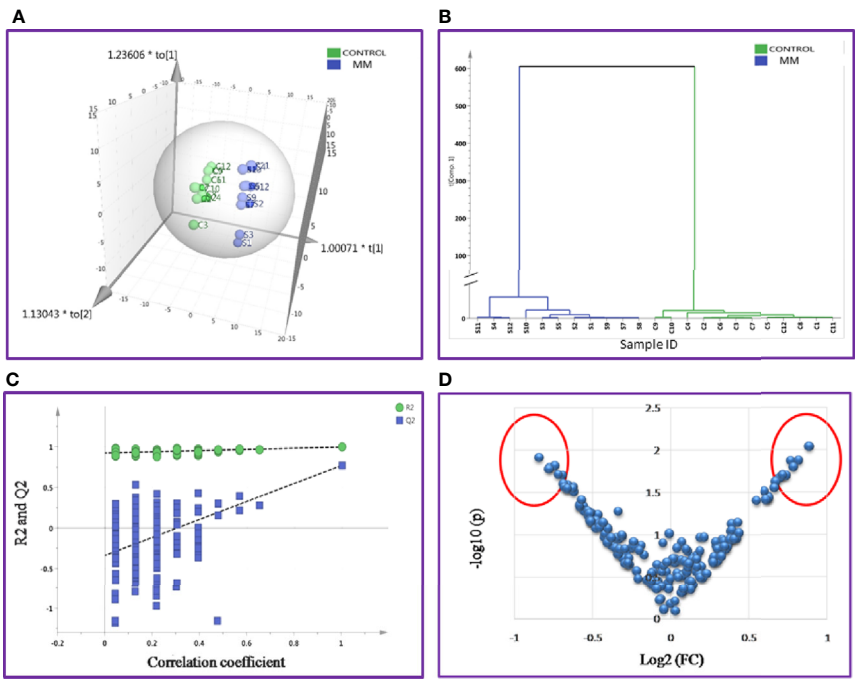


FIGURE 5 | Multivariate statistical analysis of MM serum *vs.* Control dataset using SWATH MS data. **(A)** OPLS-DA score plot for MM *vs.* Controls. **(B)** Dendrogram showing the hierarchical clustering between MM and controls based on DEPs. **(C)** Permutation test statistics of the OPLS-DA model of MM *vs.* Controls with Y-axis intercepts: R2 = (0.0, 0.924), Q2 = (0.0, −0.338), **(D)** V-Plot showing the most significant proteins in red circles. One sample in MM was removed from the SIMCA analysis as they had high degree of variation; hence the data shown is for eleven MM samples and twelve controls.

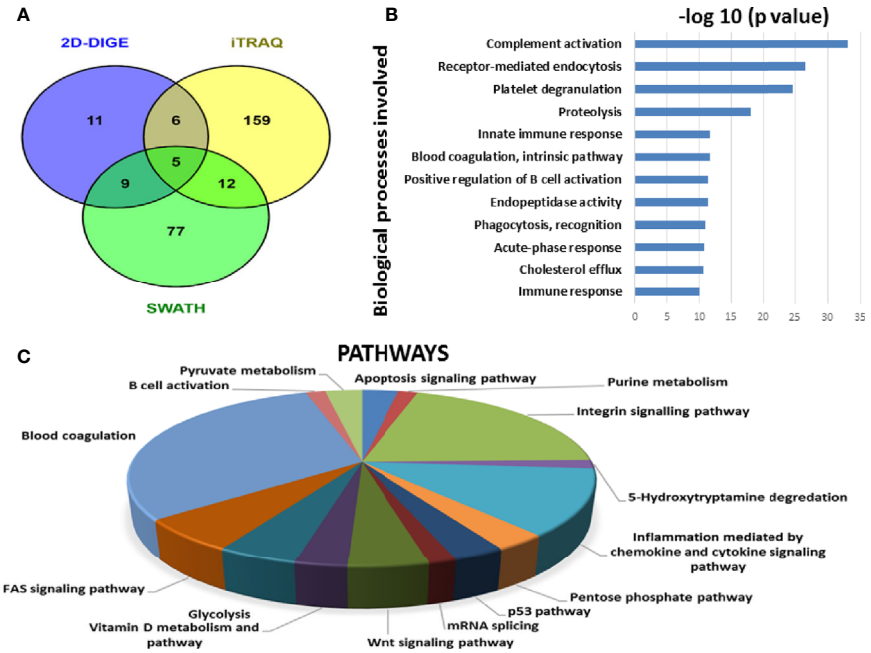


FIGURE 6 | Bioinformatic analysis of the BMIF differentially regulated protein dataset. **(A)** Representative Venn diagram for the differentially regulated proteins found through all three proteomic approaches used for bioinformatics analysis. **(B)** Biological process of the differentially regulated proteins using DAVID and **(C)** Pathways altered in the MM BMIF vs. Control BMIF dataset using PANTHER.

TABLE 1 | Partial list of differentially regulated BMIF proteins identified in MM.

Sr. No	UniProt ID	Protein Name	FC in DIGE	FC in iTRAQ	FC in SWATH	Molecular function	Biological process
1	P00738	Haptoglobin	1.78	1.51	1.55	antioxidant activity	Acute-phase response
2	P01857	Ig gamma-1 chain C region	2.49	5.84	0.45	antigen binding	B cell receptor signaling pathway
3	P01591	Ig J chain	1.69	5.70	20.27	antigen binding	adaptive immune response
4	P01871	Ig mu chain C region	0.28	6.65	15.71	antigen binding	adaptive immune response
5	P0DJI8	Serum amyloid A protein	2.38	1.89	—	chemoattractant activity	activation of MAPK activity
6	P01834	Ig kappa chain C region	—	9.49	4.98	antigen binding	B cell receptor signaling pathway
7	P01767	Ig heavy chain V-III region	—	3.13	2.34	antigen binding	complement activation
8	P00450	Ceruloplasmin	—	2.52	2.33	chaperone binding	cellular iron ion homeostasis
9	P00747	Plasminogen	1.57	—	2.29	apolipoprotein binding	blood coagulation
10	P04004	Vitronectin	2.25	—	1.63	collagen binding	cell adhesion
11	P01009	Alpha-1-antitrypsin	1.95	4.23	—	identical protein binding	acute-phase response
12	P01011	Alpha-1-antichymotrypsin	2.1	—	1.96	DNA binding	acute-phase response
13	P25311	zinc-alpha-2-glycoprotein	1.83	2.38	—	ribonuclease activity	transmembrane transport
14	P01042	Kininogen	3.94	—	2.82	endopeptidase inhibitor activity	blood coagulation, intrinsic pathway
15	P02768	Serum albumin	0.23	0.52	0.52	antioxidant activity	cellular protein metabolic process
16	P02675	Fibrinogen beta chain	0.34	0.36	—	chaperone binding	adaptive immune response
17	P80748	Ig lambda chain V-III region	—	0.67	0.54	antigen binding	complement activation
18	P01024	Complement C3		0.60	0.66	endopeptidase inhibitor activity	inflammatory response
19	P05452	Tetranectin	—	0.55	0.47	calcium ion binding	bone mineralization
20	P02786	Transferrin receptor protein 1	—	0.22	0.17	double-stranded RNA binding	cellular iron ion homeostasis
21	P02774	Vitamin D binding protein	0.25	—	0.64	actin binding	vitamin D metabolic process
22	P02787	Serotransferrin	0.38	—	0.42	ferric iron binding	actin filament organization
23	P02647	Apolipoprotein A-1	0.23	0.21	—	amyloid-beta binding	cellular protein metabolic process
24	P06396	Gelsolin	0.61	—	0.48	actin binding	actin filament capping

Bioinformatic Analysis

By employing online web applications such as PANTHER, DAVID and IPA, we tried to extract the biological information from the differentially regulated proteins identified in MM BMIF. A total of 279 statistically significant and dysregulated non-redundant proteins identified from multipronged proteomic analyses were investigated. The key biological processes found to be altered in MM were complement activation, receptor-mediated endocytosis, platelet degranulation, proteolysis, immune response, blood coagulation, positive regulation of B-cell activation, endopeptidase activity,

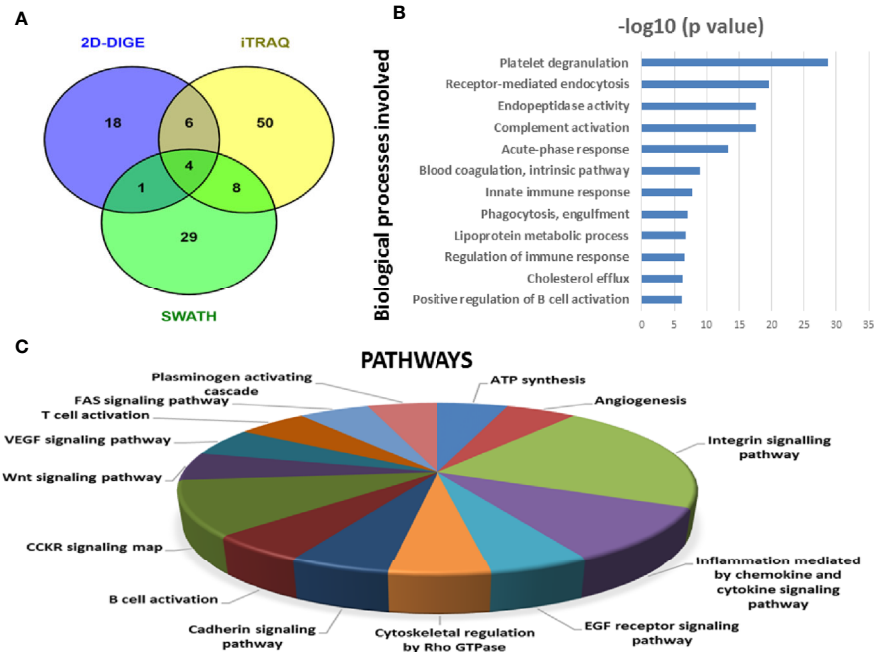


FIGURE 7 | Bioinformatic analysis of the differentially regulated protein dataset. **(A)** Representative Venn diagram for the differentially regulated proteins found through all three proteomic approaches used for bioinformatics analysis. **(B)** Biological process of the differentially regulated proteins using DAVID and **(C)** Pathways altered in the MM serum vs. Control serum dataset using PANTHER.

TABLE 2 | Partial list of differentially regulated serum proteins identified in MM.

Sr. No	UniProt ID	Protein Name	FC in DIGE	FC in iTRAQ	FC in SWATH	Molecular function	Biological process
1	P00738	Haptoglobin	1.66	1.76	2.22	Antioxidant activity	Acute-phase response
2	P01011	Alpha-1-antichymotrypsin	2.1	1.56	1.95	DNA binding	Acute-phase response
3	P01042	Kininogen 1	2.87	—	—	Cysteine-type endopeptidase	Cellular protein metabolic process
4	P0DJI8	Serum amyloid A-1 protein	—	1.95	8.25	Chemoattractant activity	Activation of mapk activity
5	P25311	Zinc-alpha-2-glycoprotein	1.83	—	—	Glycoprotein binding	Cell adhesion
6	P01009	Alpha-1-antitrypsin	1.95	—	—	Glycoprotein binding	Acute-phase response
7	P00450	Ceruloplasmin	2.16	—	—	Chaperone binding	Iron ion homeostasis
8	P02763	Alpha-1-acid glycoprotein 1	—	1.63	2.27	—	Acute-phase response
9	Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	—	1.88	1.97	Endopeptidase inhibitor activity	Hyaluronanmetabolic process
10	P01591	Ig J chain	4.38	3.30	1.66	Antigen binding	Immune response
11	P61769	Beta-2-microglobulin	—	1.73	—	Glycoprotein binding	Antibacterial humoral response
12	P43652	Afamin	—	0.38	0.61	Vitamin E binding	Vitamin transport
13	P43251	Biotinidase	—	0.38	0.57	Biotinidase activity	Biotin metabolic process
14	P29622	Kallistatin	—	0.30	0.54	Serine-type endopeptidase	Platelet degranulation
15	P05452	Tetranectin		0.51	0.5	Calcium ion binding	Bone mineralization
16	P04004	Vitronectin	0.61	0.66	—	Collagen binding	Cell adhesion
17	P02647	Apolipoprotein A1	0.25	0.53	—	Amyloid-beta binding	Adrenal gland development
18	P02787	Serotransferrin	0.38	—	—	Ferric iron binding	Iron ion homeostasis
19	P01024	Complement C3	0.43	—	—	C5L2 anaphylatoxin chemotactic receptor binding	amyloid-beta clearance
20	P02774	Vitamin D-binding protein	0.38	—	—	Actin binding	Vitamin D metabolic process
21	P02743	Serum amyloid p component	0.46	—	—	Calcium ion binding	Acute-phase response
22	P06396	Gelsolin		—	0.56	Actin binding	Actin filament capping
23	P02768	Serum albumin	0.26	—	0.14	Antioxidant activity	Bile acid and bile salt transport

phagocytosis recognition, acute phase response, cholesterol efflux, innate immune response (**Figure 6B**). Various cancer-related pathways emerged to be altered in MM such as integrin signaling pathway, apoptosis signaling pathway, P-53 pathway, Wnt signaling pathway and FAS signaling pathway (**Figure 6C**). Different molecular functions, biological processes, cellular components and proteins classes were found to be involved in MM pathogenesis (**Figure S1**). Likewise, many canonical pathways, upstream regulators and toxicology functions were identified, through IPA software, to be associated with MM BMIF samples (**Figure S2**). IPA analysis identified the connection of several upstream regulators like HNF4A, dexamethasone, lipopolysaccharide, PPARA, TGFB1 *etc*., for the differentially regulated proteins of MM BMIF.

Similarly, differentially expressed proteins in MM serum samples were also analyzed along the same lines as described above. A total of 116 statistically significant dysregulated non-redundant proteins identified from multipronged quantitative proteomic approaches were subjected to bioinformatic analysis. The majority of the key biological processes that were altered in BMIF were also found to be dysregulated in serum. These processes include platelet degranulation, endopeptidase activity, receptor-mediated endocytosis, acute phase response, complement activation, innate immune response, B cell receptor signaling and phagocytosis (**Figure 7B**). Moreover, various cancer-related pathways emerged to be altered in MM such as angiogenesis, integrin signaling pathway, Wnt signaling pathway and EGF receptor signaling pathway (**Figure 7C**). Furthermore, different molecular functions, biological processes, cellular components and protein classes were also found to be involved in MM pathogenesis (**Figure S3**). In addition, IPA based analysis revealed many canonical pathways, upstream regulators and toxicology functions to be associated with MM serum samples (**Figure S4**). IPA analysis identified the involvement of several upstream regulators like IL6, nitrofurantoin, HNF1A, HNF4A, *etc*., for the differentially regulated proteins of MM serum.

Verification Experiments

To verify the differentially expressed proteins found in the MM BMIF discovery phase of our study, we performed the verification experiments using two approaches *viz*. the western blotting and the mass spectrometry based SRM assay. Western blotting was performed for Kininogen 1, alpha-1-antitrypsin, vitronectin, gelsolin, apolipoprotein A1, and transferrin. The immunoblot data indicated a consistent pattern of differential expression as observed in discovery phase data (**Figure S5**). We also performed the SRM verification for some of the proteins that had not been verified through the western blotting due to the unavailability of antibodies in our lab and to adopt this cost-effective SRM verification strategy. Proteins such as ceruloplasmin, haptoglobin, apolipoprotein A-IV, alpha-1-acid glycoprotein 1, beta-2-glycoprotein, vitronectin, plasminogen, serum amyloid p-component, complement C3, transferrin, apolipoproteinA1, and fibrinogen alpha chain were verified by this approach (**Figure S6**). The results obtained in the SRM based verification were in good agreement with the discovery phase data.

Similarly, for selected proteins that are differentially expressed in MM serum, immunoblotting was performed for their verification. These proteins include haptoglobin, kininogen 1, zinc alpha-2-glycoprotein, gelsolin, apolipoprotein A1, and transferrin (**Figure S7**). Moreover, the proteins which are mentioned above along with few other proteins such as alpha-1-antitrypsin, ceruloplasmin, plasminogen, serum amyloid A1, complement C3, vitamin D binding protein, serum amyloid P component and alpha-1-antichymotrypsin were also verified by SRM based approach (**Figure S8**). The findings of immunoblot and SRM based assays were consistent with our discovery phase observations concerning either up-regulation or down-regulation of the above mentioned proteins.

Validation Experiments

The BMIF, as well as serum proteins that showed statistically significant differential abundance in the discovery phase and verified for their expression by immunoblotting or SRM based approach were further validated in an external cohort of MM patients' serum samples by ELISA (MM = 36, Control = 36). ELISA based validation experiments were performed for proteins *viz*., haptoglobin, kininogen 1, apolipoprotein A1, transferrin and albumin. Among these proteins, haptoglobin and kininogen 1 were found to be up-regulated, whereas apolipoprotein A1, transferrin and albumin proteins were down-regulated in MM samples. Serum albumin in MM patients was found to be at 26.45 mg/ml ± 8.14 mg/ml, while in healthy controls it was at 42.80 mg/ml ± 10.13 mg/ml. Serum apolipoprotein A1 in MM patients was at a concentration of 105.97 mg/dl ± 38.80 mg/dl whereas in healthy controls it was at 152.05 mg/dl ± 30.61 mg/dl concentration. The concentration range of serum kininogen 1 in MM patients was 173.95 μg/dl ± 26.60 μg/dl whereas in healthy controls it was 139.38 μg/dl ± 28.98 μg/dl. MM patients had a concentration of 145.28 mg/dl ± 45.03 mg/dl for serum haptoglobin, which was found to be decreased in healthy controls at 110.86 mg/dL ± 30.99 mg/dL. Further, the concentration of serum transferrin in MM patients was at 1.76 g/L ± 0.48 g/L whereas in healthy controls it was found to be 2.64 g/L± 0.69 g/L. The concentration range of these proteins in healthy controls as well as in MM samples is showed in **Figure 8**. Further, these proteins were investigated for their ability to discriminate the MM patients from healthy controls by building the Receiver Operating Characteristic (ROC) curve analysis (**Figure S9**).

DISCUSSION

MM is the second most common hematological malignancy associated with plasma cell deformation and ultimately leads to bone marrow resorption. Currently, MM diagnosis primarily depends on the protein markers like serum albumin and B2M which are lesser specific at the early stages of this disease. Therefore, finding the novel and sensitive diagnostic and therapeutic markers for MM has a tremendous potential to improve patient survival. BMIF is the proximal biofluid for hematological malignancies and could serve as a potential

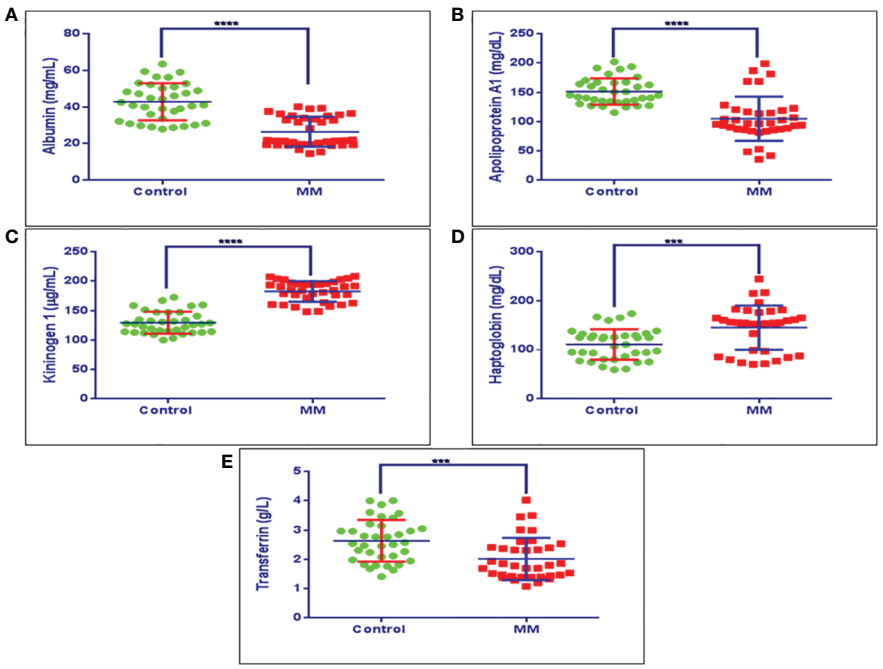


FIGURE 8 | Validation experiments for some of the candidate proteins using ELISA. **(A)** Concentration ranges of serum albumin in healthy controls (42.80 mg/ml ± 10.13 mg/ml), MM patients (26.45 mg/ml ± 8.14 mg/ml). **(B)** Concentration ranges of serum apolipoprotein a1 in healthy controls (152.05 mg/dl ± 30.61 mg/dl), MM patients (105.97 mg/dl ± 38.80 mg/dl). **(C)** Concentration ranges of serum kininogen 1 in healthy controls (139.38 µg/dl ± 28.98 µg/dl), MM patients (173.95 µg/dl ± 26.60 µg/dl), **(D)** Concentration ranges of serum haptoglobin in healthy controls (110.86 mg/dl ± 30.99 mg/dl), MM patients (145.28 mg/dl ± 45.03 mg/dl). **(E)** Concentration ranges of serum transferrin in healthy controls (2.64 g/L ± 0.69 g/L), MM patients (1.76 g/L ± 0.48 g/L). Healthy control serum samples (n = 36), MM patients serum samples (n = 36). ****p-value ≤ 0.0001, ***p-value ≤ 0.001.

source for identifying diagnostic, prognostic and therapeutic markers for MM. But unfortunately, bone marrow aspiration is an invasive and tedious procedure that puts patients through considerable anxiety and stress. Hence, to overcome this problem, for the first time, we identified and verified a set of differentially abundant proteins in MM BMIF and serum samples using multipronged quantitative proteomic approaches. Further, we selected a set of common proteins that showed a similar pattern of differential abundance in both MM BMIF as well as serum. Finally, we validated a panel of potential candidate biomarkers in a large cohort of serum samples that is associated with MM.

The bioinformatics analysis of the significant differentially abundant proteins found to be involved in many biological processes, some of which are elaborated henceforth in context to the proteins identified in this study. Platelet degranulation has been reported as a major biological process being altered in various cancers due to differentially expressed proteins (37, 38). In this study, platelet degranulation proteins such as complement C3, serum albumin, alpha-2-HS-glycoprotein, transferrin, extracellular matrix protein 1 and tetranectin were seen to be down-regulated in both MM BMIF as well as serum. A report from Gay et al. infers that platelet activation followed by degranulation serves as an important phenomenon in cancer (37). They have suggested that platelets in the circulatory system help the cancer cells to evade the immune response and enhance their malignant potential. We have found some proteins involved in this process to be down-regulated which is in agreement with the earlier report by Gay and co-authors (37). Albumin is one of the well-established biomarkers and is used in the international staging for the MM, the lower levels of albumin is associated with greater disease severity (39, 40). It regulates the platelet degranulation *via* its down-regulation, which is also evident from a similar expression pattern found in our study.

MM is a plasma cell cancer and these cells are derived from B cells. Hence, any alterations in the B-cell pathways such as the positive regulation of B cell activation and B cell receptor (BCR) signaling pathway, as observed in this study, could play an important role in MM development and progression. Interestingly, differentially expressed proteins *viz*., Ig kappa chain C region, Ig mu chain C region, Ig alpha-1 chain C region and Ig heavy chain V-III region involved in these pathways depicted an up-regulated expression pattern in both biological fluids. BCR signaling plays a key role in the maintenance and development of B cells. The pathways altered downstream of BCR signaling would lead to the proliferation and survival of the B cells. The differentially expressed proteins associated with the BCR signaling pathway found in our study were positively modulated, which could lead to proliferative signals leading to the development of MM as explained by

Choi & Kipps (41). Various lymphomas utilize BCR signaling as a key oncogenic pathway to promote the proliferation and survival of B cells (42–44).

Complement activation is an important cause for inflammation and a series of experiments have proved that it could lead to tumor progression (45). Complement component C3-deficient mice were found to escape chemically induced carcinogenesis in different tissues due to reduced inflammation (46). They also identified the long pentraxin (a humoral component of innate immunity) which acts as a negative regulator of complement activation and cancer-related inflammation and further showed that PTX3 knockout mice were prone to carcinogenesis. In our study, we observed that the majority of proteins related to complement activation pathways such as Ig kappa chain C region, Ig J chain, Complement C4-B, Ig gamma-1 chain C region and Protein IGKV3-11 showed increased expression pattern.

Previous studies have reported that higher levels of intracellular cholesterol positively affect cancer cell proliferation and migration (47, 48). The positive regulation between the elevation of cellular cholesterol and tumorigenesis mechanism is not clearly reported and needs to be explored. Our study identified several proteins involved in cholesterol efflux such as apolipoprotein A-4, apolipoprotein A-1, apolipoprotein C-II, apolipoprotein M, apolipoprotein L1, apolipoprotein E, apolipoprotein D and apolipoprotein C1 to be down-regulated in MM BMIF and serum. Similarly, the lipoprotein metabolomic process was also altered and most of the proteins were the same as those found in cholesterol efflux.

Some of the acute phase proteins have been related to distinct cancers and also linked to their malignancy stages. Acute-phase proteins like alpha-1-acid glycoprotein 1, inter alpha- trypsin, serum amyloid A protein, alpha-1-antitrypsin, serum amyloid A-4 protein, C-reactive protein (CRP) showed up-regulation in this study. Previously, few reports demonstrated CRP levels to be elevated in squamous cell carcinoma of the esophagus and in adenocarcinoma and revealed that an increase in CRP levels correlated with tumor growth and metastasis (49–51). Ilhan et al. and Chan et al. demonstrated that CRP and serum amyeloid A-1 proteins were elevated in gastric cancer and SAA was also found to be elevated in recurrent gastric cancer patients (52, 53).

During the tumor progression, various innate immunity components are activated to minimize the inflammation caused by cancer (54, 55). These activated innate immunity components further positively affect the adaptive immune responses and help in targeting the cancer cells (54, 55). In our study, innate immunity-related proteins such as Ig J chain, Ig lambda-like polypeptide 5, Ig alpha-1 chain C region, serum amyloid A-1 protein, Ig kappa chain C region, beta-2-microglobulin and complement C4-B showed up-regulation. The B2M protein level is elevated in MM and is an approved diagnostic marker for MM detection that was found to be up-regulated in our MM serum samples as well. Various other biological processes such as endopeptidase activity, receptor-mediated endocytosis, blood coagulation, phagocytosis and immune response were also found to be altered due to the differentially expressed proteins suggesting their potential role in the MM.

Cancer-associated pathways such as integrin signaling pathway, inflammation mediated by chemokine and cytokine signaling pathway, Wnt signaling pathway and FAS signaling pathway appeared to be altered in both MM BMIF as well as serum. The integrin signaling pathway plays an important role in cancer growth, metastasis, and therapy resistance in the tumor cells as well as in stromal cells. Integrins involve in the interaction of cells with their local environment and translate the external chemical response into a concerted intracellular response (56). Chemokine signaling pathways regulate the immune responses, angiogenesis, epithelial cell growth and survival. Chemokines are also critical for cancer progression and play an important role in the tumor microenvironment (57). Likewise, cytokine inflammation signaling pathways linked to chronic diseases such as obesity, heart diseases and cancer (58). Similarly, various research studies reveal that the Fas/FasL signaling and Wnt signaling pathway plays a crucial role in the impairment of cancer cells and leads to apoptosis resistance and tumor progression (59, 60). Hence, alterations of these pathways due to the dysreulated proteins could have an important role in the MM cancer progression.

Further, we performed the verification experiments using two approaches *viz*., the western blotting and the mass spectrometry based SRM assay to verify the proteins differentially regulated in the discovery phase. A panel of 15 Proteins *viz*., kininogen 1, alpha-1-antitrypsin, vitronectin, gelsolin, transferrin, ceruloplasmin, haptoglobin, apolipoprotein A-IV, alpha-1-acid glycoprotein 1, beta-2-glycoprotein, plasminogen, serum amyloid p-component, complement C3, apolipoprotein A1, and fibrinogen alpha chain were verified in MM BMIF. Similarly, a panel of 14 proteins *viz*. haptoglobin, kininogen 1, alpha-1-antitrypsin, zinc-2-alpha glycoprotein, gelsolin, apolipoprotein A1, transferrin, ceruloplasmin, plasminogen, serum amyloid A1, complement C3, vitamin D binding protein, serum amyloid P component and alpha-1-antichymotrypsin were verified in an external cohort of serum samples. Finally, we validated a panel of five of the verified proteins such as haptoglobin, kininogen 1, transferrin and apolipoprotein A1 along with known MM biomarker, albumin in a large cohort of serum samples using ELISA. This panel of proteins could be useful in the diagnosis and understanding of the pathophysiology of MM. The relevance of each protein in the potential biosignature panel of MM identified in this study is discussed henceforth.

Haptoglobin is an acute phase alpha glycoprotein comprised of two alpha and two beta subunits and tasked with hemoglobin clearance upon erythrocyte lysis. Haptoglobin is also known to carry out several other functions including antioxidant, anti-inflammatory and immune response regulation (61). Moreover, a recent study has shown that elevated levels of serum haptoglobin regarded as a diagnostic and prognostic marker of non-small-cell lung carcinoma (NSCLC). The study indicated that higher levels of haptoglobin have positively correlated with disease progression as well as distant organ metastasis in NSCLC

(62). Furthermore, in another study, haptoglobin alpha subunit was found to be highly overexpressed in serum samples of ovarian cancer patients as compared to controls. The authors of the study projected haptoglobin alpha subunit as a complementary diagnostic marker along with established CA 125 protein (63). Similarly, in our study, we also observed consistent up-regulation of haptoglobin in MM BMIF and serum samples as compared to controls. Furthermore, the up-regulation of haptoglobin was observed in all three approaches *viz*., 2D-DIGE, iTRAQ, SWATH in MM BMIF and serum samples suggesting a potential candidate for MM. Overall, the MM serum and BMIF haptoglobin expression profiles were in good agreement with previous reports.

Alternate splicing of kininogen 1 transcript produces high molecular weight kininogen (HMWK) and low molecular weight kininogen (LMWK). The HMWK is involved in blood coagulation as well as the proteolytic release of bradykinin peptide (64). Bradykinin plays an important role in several physiological activities such as antimicrobial, smooth muscle contraction and release of inflammatory mediators like prostaglandins (65). Moreover, bradykinin is also known to interact with EGFR and stimulate downstream signaling pathways to enhance cell invasion as well as promote angiogenesis through elevated VEGF expression in glioma (66, 67). Furthermore, higher expression of kininogen 1 is recently reported as a serum biomarker of advanced colorectal adenoma, colorectal cancer and prognostic marker of oral cancer (68, 69). Interestingly, in our study, we also noticed elevated levels of kininogen 1 in BMIF as well as serum samples of MM patients as compared to the controls in discovery as well as validation cohorts. The findings were consistent with earlier reports of higher expression of kininogen 1 under the malignant condition and could be projected as a potential candidate biomarker for MM.

Transferrin is an iron transport protein present in blood serum. The total concentration of blood transferrin usually shown as a total iron-binding capacity of serum (TIBC). Lower TIBC leads to hemolytic anemia in which red blood cells destroyed very early stages (70). Interestingly, it is well known that anemia is a common characteristic feature of MM (71). Furthermore, it is also reported that low TIBC was significantly associated with lower disease-free survival as well as the overall survival rate of gastric cancer patients (72). Likewise, Hodgkin's lymphoma is also reported to be associated with a low serum iron concentration and reduced transferrin saturation (73). In contrast to the above mentioned studies, few studies have also shown that transferrin could be a positive regulator for the cancer cells proliferation by acting as an autocrine regulator (74). In our study, we observed reduced levels of transferrin in MM BMIF and serum samples as compared to controls. Since, transferrin is associated with anemia which is a well known characteristic feature of MM, it could be a potential protein marker for MM.

Apolipoprotein A-1 is a multifunctional component of high-density lipoprotein (HDL) involved in inflammation and immune response regulation apart from cholesterol trafficking. Recently, many studies have observed that apolipoprotein A-1 levels are altered upon cancer development and indicated that it could serve as a useful marker for diagnosis, prognosis as well as risk stratification of cancer patients (75). It's been also reported that reduced levels of apolipoprotein A-1 have been observed in several cancers including adenocarcinoma of the gastrointestinal tract, lung and breast adenocarcinoma, early-stage ovarian cancer, cervical cancer and lymphoblastic leukemia (76–82). Furthermore, it's also documented that apolipoprotein A-1 levels were significantly lowered upon the metastatic recurrence of the liver, breast, endometrial and cervical cancer (83–86). Moreover, a prominent association of apolipoprotein A-1 was also revealed in the prognosis of several cancers (87). In our study, we also noticed reduced levels of apolipoprotein A-1 in MM BMIF and serum samples as compared to controls. The findings were consistent with earlier reports but, further studies are needed to probe the precise mechanism of apolipoprotein A-1 role in MM manifestation.

Albumin is a globular protein synthesized in the liver and performs the primary function of regulating the oncotic pressure of blood apart from acting as a carrier protein for various hormones, vitamins, and enzymes. Albumin constitutes over 60% of total blood plasma protein and its differential abundance levels have been intrinsically associated with several ailments. Importantly, in one of the study, hypoalbuminemia, a condition of reduced albumin levels have been detected as a marker of the advanced stage as well as higher cancer burden in MM patients (88). Authors have observed that the albumin levels ≤29.0 g/L indicate advanced disease stage of MM. Furthermore, in another study, reduced levels of serum albumin (<3.5 g/dl) projected as a significant prognosis factor in symptomatic MM (39). In our study, we also detected lower levels of albumin in MM BMIF as well as serum samples as compared to the controls in all three proteomic approaches. The findings of our study support the already confirmed fact that decreased abundance levels of albumin is associated with MM disease manifestation and hence, could be employed as an important member of the panel of proteins identified as a diagnostic marker in our study that could be translated to future clinical settings with higher confidence.

CONCLUSIONS

MM is a disease that is primarily diagnosed in the advanced stages where the medical interventions for the treatment are very limited. Early diagnosis of MM can be crucial as the chances of better disease management by the clinicians can increase the overall survival expectancy of the MM patients. Protein-based makers of MM can comply with finding the biosignature which could be developed as early predictor of this disease. Our study design was to identify a biosignature panel of proteins from the BMIF and serum samples of MM patients and cross-verified its abundance in an independent cohort of serum samples of MM patients as the serum is an easily obtained diagnostic biofluid. Using multipronged proteomic approaches, we identified 279 and 116 non-redundant

differentially expressed proteins in MM BMIF and serum respectively. A verification phase of experiments in an external cohort of BMIF and serum samples confirmed a set of 15 and 14 proteins respectively. Finally, a combined panel of four common proteins namely haptoglobin, kininogen 1, transferrin and apolipoprotein A1 along with albumin (an established biomarker for MM) were validated in a fresh cohort of serum samples and could be better and minimally invasive diagnostic, prognostic markers for MM. However, validation in a larger cohort of MM patients can further help to investigate the practical potentiality of these proteins as early diagnostic and prognostic markers. We believe that this panel of proteins could help in future MM disease management and thereby improving the survival expectancy of MM patients.

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: Proteome Xchange Consortium *via* the PRIDE partner repository (89) having identifier PXD008927.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Armed Forces Medical College, Pune and National Centre for Cell Science, Pune. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceived the study: VC, TC, SR. Designed the study: VC, KT, TC, SS, VS, MS, SR. Performed the experiments: VC, FR. Compiled the data: VC, RT, KT, FR, DM, SR. Analyzed the proteomics data and performed the statistical and bioinformatics analysis: VC, RT, KT, FR, DM, SR. Drafted the manuscript: VC, RT, KT, TC, SS, FR, DM, MS, SR. Provided clinical samples: TC, SS, VS. Provided chemicals and reagents: SR. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by NCCS intramural funding and the Department of Biotechnology, Govt. of India, India (grant no. BT/PR10855/BRB/10/1330/2014).

ACKNOWLEDGMENTS

The authors are grateful to all the volunteers who participated in this study. VC, RT, and KT acknowledge the Council of Scientific & Industrial Research (CSIR) for fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.566804/full#supplementary-material

REFERENCES

1. Kyle RA, Rajkumar SV. Multiple myeloma. *Blood* (2008) 111(6):2962–72. doi: 10.1182/blood-2007-10-078022
2. Landgren O, Weiss BM. Patterns of monoclonal gammopathy of undetermined significance and multiple myeloma in various ethnic/racial groups: support for genetic factors in pathogenesis. *Leukemia* (2009) 23 (10):1691–7. doi: 10.1038/leu.2009.134
3. Hideshima T, Bergsagel PL, Kuehl WM, Anderson KC. Advances in biology of multiple myeloma: clinical applications. *Blood* (2004) 104(3):607–18. doi: 10.1182/blood-2004-01-0037
4. Palumbo A, Anderson K. Multiple Myeloma. *N Engl J Med* (2011) 364 (11):1046–60. doi: 10.1056/NEJMra1011442
5. Scudla V, Zemanova M, Minarik J, Bacovsky J, Ordeltova M, Indrak K, et al. International prognostic index (IPI)–a critical comparison with five multiple myeloma staging systems in the group of 270 patients treated by conventional chemotherapy. *Neoplasma* (2006) 53(4):277–84.
6. Tichy M, Maisnar V, Palicka V, Friedecky B, Vavrova J, Novotna H, et al. International Staging System required standardization of biochemical laboratory testing in multiple myeloma. *Neoplasma* (2006) 53(6):492–4.
7. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J, et al. International staging system for multiple myeloma. *J Clin Oncol* (2005) 23 (15):3412–20. doi: 10.1200/JCO.2005.04.242
8. Avet-Loiseau H, Durie BG, Cavo M, Attal M, Gutierrez N, Haessler J, et al. Combining fluorescent in situ hybridization data with ISS staging improves risk assessment in myeloma: an International Myeloma Working Group collaborative project. *Leukemia* (2013) 27(3):711–7. doi: 10.1038/leu.2012.282
9. Aebersold R, Goodlett DR. Mass Spectrometry in Proteomics. *Chem Rev* (2001) 101(2):269–96. doi: 10.1021/cr990076h
10. Alaoui-Jamali MA, Xu Y-j. Proteomic technology for biomarker profiling in cancer: an update. *J Zhejiang Univ Sci B* (2006) 7(6):411–20. doi: 10.1631/jzus.2006.B0411
11. Hristova VA, Chan DW. Cancer biomarker discovery and translation: proteomics and beyond. *Expert Rev Proteom* (2019) 16(2):93–103. doi: 10.1080/14789450.2019.1559062
12. Kellie JF, Sikorski TW, An B, Chen Z, Moghieb AH, Busz MG, et al. A new era for proteomics. *Future Sci* (2019) 11(19):1731–5. doi: 10.4155/bio-2019-0191
13. Macklin A, Khan S, Kislinger T. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Bioanalysis* (2020) 17:1–25. doi: 10.1186/s12014-020-09283-w
14. Chen J, Zheng N. Accelerating protein biomarker discovery and translation from proteomics research for clinical utility. *Bioanalysis* (2020) 12(20):1469–81. doi: 10.4155/bio-2020-0198
15. Ge F, Bi LJ, Tao SC, Xu XD, Zhang ZP, Kitazato K, et al. Proteomic analysis of multiple myeloma: current status and future perspectives. *Proteomics Clin Appl* (2011) 5(1-2):30–7. doi: 10.1002/prca.201000044
16. Micallef J, Dharsee M, Chen J, Ackloo S, Evans K, Qiu L, et al. Applying mass spectrometry based proteomic technology to advance the understanding of multiple myeloma. *J Hematol Oncol* (2010) 3:13. doi: 10.1186/1756-8722-3-13
17. Ge F, Xiao CL, Yin XF, Lu CH, Zeng HL, He QY. Phosphoproteomic analysis of primary human multiple myeloma cells. *J Proteom* (2010) 73(7):1381–90. doi: 10.1016/j.jprot.2010.03.004
18. Ge F, Li WL, Bi LJ, Tao SC, Zhang ZP, Zhang XE. Identification of novel 14-3-3zeta interacting proteins by quantitative immunoprecipitation combined

with knockdown (QUICK). *J Proteome Res* (2010) 9(11):5848–58. doi: 10.1021/pr100616g
19. Xiao CL, Zhang ZP, Xiong S, Lu CH, Wei HP, Zeng HL, et al. Comparative proteomic analysis to discover potential therapeutic targets in human multiple myeloma. *Proteomics Clin Appl* (2009) 3(11):1348–60. doi: 10.1002/prca.200900068
20. Bai J, Yang Y, Wang J, Zhang L, Wang F, He A. Variability of serum novel serum peptide biomarkers correlates with the disease states of multiple myeloma. *Clin Proteom* (2019) 16(1):17. doi: 10.1186/s12014-019-9238-0
21. Ma TZ, Piao Z, Jin SY, Kwak YG. Differential expression of serum proteins in multiple myeloma. *Exp Ther Med* (2019) 17(1):649–56. doi: 10.3892/etm.2018.7010
22. Zhang HT, Tian EB, Chen YL, Deng HT, Wang QT. Proteomic analysis for finding serum pathogenic factors and potential biomarkers in multiple myeloma. *Chin Med J (Engl)* (2015) 128(8):1108–13. doi: 10.4103/0366-6999.155112
23. Wang QT, Li YZ, Liang YF, Hu CJ, Zhai YH, Zhao GF, et al. Construction of a multiple myeloma diagnostic model by magnetic bead-based MALDI-TOF mass spectrometry of serum and pattern recognition software. *Anatomical Record (Hoboken NJ 2007)* (2009) 292(4):604–10. doi: 10.1002/ar.20871
24. Chun-Hua L, Feng G, Zhi L, Rong L, Chuan-Le X, Hui-Lan Z, et al. Detection of Abundant Proteins in Multiple Myeloma Cells by Proteomics. *J Proteomics Bioinform* (2010) - 3(- 1):005–9. doi: 10.4172/jpb.1000115
25. Rajpal R, Dowling P, Meiller J, Clarke C, Murphy WG, O'Connor R, et al. A novel panel of protein biomarkers for predicting response to thalidomide-based therapy in newly diagnosed multiple myeloma patients. *Proteomics* (2011) 11(8):1391–402. doi: 10.1002/pmic.201000471
26. Ge F, Zhang L, Tao SC, Kitazato K, Zhang ZP, Zhang XE, et al. Quantitative proteomic analysis of tumor reversion in multiple myeloma cells. *J Proteome Res* (2011) 10(2):845–55. doi: 10.1021/pr100992e
27. Cumova J, Potacova A, Zdrahal Z, Hajek R. Proteomic analysis in multiple myeloma research. *Mol Biotechnol* (2011) 47(1):83–93. doi: 10.1007/s12033-010-9326-x
28. Jia H, Ge F, Lu X, Zeng H, Li L, Chen Z, et al. Proteomics of apoptosis of multiple myeloma cells induced by proteasome inhibitor PS-341. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* (2010) 35(8):784–91. doi: 10.3969/j.issn.1672-7347.2010.08.003
29. Ge F, Xiao CL, Bi LJ, Tao SC, Xiong S, Yin XF, et al. Quantitative phosphoproteomics of proteasome inhibition in multiple myeloma cells. *PloS One* (2010) 5(9):e13095. doi: 10.1371/journal.pone.0013095
30. Gjertsen BT, Wiig H. Investigation of therapy resistance mechanisms in myeloid leukemia by protein profiling of bone marrow extracellular fluid. *Expert Rev Proteom* (2012) 9(6):595–8. doi: 10.1586/epr.12.55
31. Gajbhiye A, Dabhi R, Taunk K, Vannuruswamy G, RoyChoudhury S, Adhav R, et al. Urinary proteome alterations in HER2 enriched breast cancer revealed by multipronged quantitative proteomics. *Proteomics* (2016) 16(17):2403–18. doi: 10.1002/pmic.201600015
32. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* (2006) 1(6):2856–60. doi: 10.1038/nprot.2006.468
33. Korwar AM, Vannuruswamy G, Jagadeeshaprasad MG, Jayaramaiah RH, Bhat S, Regin BS, et al. Development of Diagnostic Fragment Ion Library for Glycated Peptides of Human Serum Albumin: Targeted Quantification in Prediabetic, Diabetic, and Microalbuminuria Plasma by Parallel Reaction Monitoring, SWATH, and MSE. *Mol Cell Proteom* (2015) 14(8):2150–9. doi: 10.1074/mcp.M115.050518
34. Pendharkar N, Gajbhiye A, Taunk K, RoyChoudhury S, Dhali S, Seal S, et al. Quantitative tissue proteomic investigation of invasive ductal carcinoma of breast with luminal B HER2 positive and HER2 enriched subtypes towards potential diagnostic and therapeutic biomarkers. *J Proteomics* (2016) 132:112–30. doi: 10.1016/j.jprot.2015.11.024
35. Kume H, Muraoka S, Kuga T, Adachi J, Narumi R, Watanabe S, et al. Discovery of colorectal cancer biomarker candidates by membrane proteomic analysis and subsequent verification using selected reaction monitoring (SRM) and tissue microarray (TMA) analysis. *Mol Cell Proteomics* (2014) 13(6):1471–84. doi: 10.1074/mcp.M113.037093
36. Maclean B, Tomazela DM, Abbatiello SE, Zhang S, Whiteaker JR, Paulovich AG, et al. Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. *Analyt Chem* (2010) 82(24):10116–24. doi: 10.1021/ac102179j
37. Gay LJ, Felding-Habermann B. Contribution of platelets to tumour metastasis. *Nat Rev Cancer* (2011) 11(2):123–34. doi: 10.1038/nrc3004
38. Egan K, Crowley D, Smyth P, O'Toole S, Spillane C, Martin C, et al. Platelet adhesion and degranulation induce pro-survival and pro-angiogenic signalling in ovarian cancer cells. *PloS One* (2011) 6(10):e26125. doi: 10.1371/journal.pone.0026125
39. Kim JE, Yoo C, Lee DH, Kim S-W, Lee J-S, Suh C. Serum albumin level is a significant prognostic factor reflecting disease severity in symptomatic multiple myeloma. *Ann Hematol* (2010) 89(4):391–7. doi: 10.1007/s00277-009-0841-4
40. Chen YH, Magalhaes MC. Hypoalbuminemia in patients with multiple myeloma. *Arch Intern Med* (1990) 150(3):605–10. doi: 10.1001/archinte.150.3.605
41. Choi MY, Kipps TJ. Inhibitors of B-cell receptor signaling for patients with B-cell malignancies. *Cancer J* (2012) 18(5):404–10. doi: 10.1097/PPO.0b013e31826c5810
42. Yan Q, Huang Y, Watkins AJ, Kocialkowski S, Zeng N, Hamoudi RA, et al. BCR and TLR signaling pathways are recurrently targeted by genetic changes in splenic marginal zone lymphomas. *Haematologica* (2012) 97(4):595–8. doi: 10.3324/haematol.2011.054080
43. Pighi C, Gu TL, Dalai I, Barbi S, Parolini C, Bertolaso A, et al. Phospho-proteomic analysis of mantle cell lymphoma cells suggests a pro-survival role of B-cell receptor signaling. *Cell Oncol (Dordr)* (2011) 34(2):141–53. doi: 10.1007/s13402-011-0019-7
44. Herishanu Y, Perez-Galan P, Liu D, Biancotto A, Pittaluga S, Vire B, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* (2011) 117(2):563–74. doi: 10.1182/blood-2010-05-284984
45. Coussens LM, Werb Z. Inflammation and cancer. *Nature* (2002) 420 (6917):860–7. doi: 10.1038/nature01322
46. Bonavita E, Gentile S, Rubino M, Maina V, Papait R, Kunderfranco P, et al. PTX3 is an extrinsic oncosuppressor regulating complement-dependent inflammation in cancer. *Cell* (2015) 160(4):700–14. doi: 10.1016/j.cell.2015.01.004
47. Llaverias G, Danilo C, Mercier I, Daumer K, Capozza F, Williams TM, et al. Role of cholesterol in the development and progression of breast cancer. *Am J Pathol* (2011) 178(1):402–12. doi: 10.1016/j.ajpath.2010.11.005
48. Notarnicola M, Messa C, Pricci M, Guerra V, Altomare DF, Montemurro S, et al. Up-regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity in left-sided human colon cancer. *Anticancer Res* (2004) 24(6):3837–42.
49. Wang CY, Hsieh MJ, Chiu YC, Li SH, Huang HW, Fang FM, et al. Higher serum C-reactive protein concentration and hypoalbuminemia are poor prognostic indicators in patients with esophageal cancer undergoing radiotherapy. *Radiother Oncol* (2009) 92(2):270–5. doi: 10.1016/j.radonc.2009.01.002
50. Gockel I, Dirksen K, Messow CM, Junginger T. Significance of preoperative C-reactive protein as a parameter of the perioperative course and long-term prognosis in squamous cell carcinoma and adenocarcinoma of the oesophagus. *World J Gastroenterol WJG* (2006) 12(23):3746–50. doi: 10.3748/wjg.v12.i23.3746
51. Nozoe T, Saeki H, Sugimachi K. Significance of preoperative elevation of serum C-reactive protein as an indicator of prognosis in esophageal carcinoma. *Am J Surg* (2001) 182(2):197–201. doi: 10.1016/S0002-9610(01)00684-5
52. Chan DC, Chen CJ, Chu HC, Chang WK, Yu JC, Chen YJ, et al. Evaluation of serum amyloid A as a biomarker for gastric cancer. *Ann Surg Oncol* (2007) 14 (1):84–93. doi: 10.1245/s10434-006-9091-z
53. Ilhan N, Ilhan N, Ilhan Y, Akbulut H, Kucuksu M. C-reactive protein, procalcitonin, interleukin-6, vascular endothelial growth factor and oxidative metabolites in diagnosis of infection and staging in patients with gastric cancer. *World J Gastroenterol WJG* (2004) 10(8):1115–20. doi: 10.3748/wjg.v10.i8.1115
54. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* (2013) 39(1):1–10. doi: 10.1016/j.immuni.2013.07.012

55. Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoediting. *Nat Rev Immunol* (2006) 6(11):836–48. doi: 10.1038/nri1961
56. Danen EH. Integrin signaling as a cancer drug target. *ISRN Cell Biol* (2013) 2013. doi: 10.1155/2013/135164
57. Hembruff SL, Cheng N. Chemokine signaling in cancer: Implications on the tumor microenvironment and therapeutic targeting. *Cancer Ther* (2009) 7 (A):254–67.
58. Lin W-W, Karin M. A cytokine-mediated link between innate immunity, inflammation, and cancer. *J Clin Invest* (2007) 117(5):1175–83. doi: 10.1172/JCI31537
59. Villa-Morales M, Fernández-Piqueras J. Targeting the Fas/FasL signaling pathway in cancer therapy. *Expert Opin Ther Targ* (2012) 16(1):85–101. doi: 10.1517/14728222.2011.628937
60. Klaus A, Birchmeier W. Wnt signalling and its impact on development and cancer. *Nat Rev Cancer* (2008) 8(5):387–98. doi: 10.1038/nrc2389
61. MacKellar M, Vigerust DJ. Role of haptoglobin in health and disease: a focus on diabetes. *Clin Diabetes* (2016) 34(3):148–57. doi: 10.2337/diaclin.34.3.148
62. Lu J, Wang Y, Yan M, Feng P, Yuan L, Cai Y, et al. High serum haptoglobin level is associated with tumor progression and predicts poor prognosis in non-small cell lung cancer. *Oncotarget* (2016) 7(27):41758. doi: 10.18632/oncotarget.9676
63. Ye B, Cramer DW, Skates SJ, Gygi SP, Pratomo V, Fu L, et al. Haptoglobin-α subunit as potential serum biomarker in ovarian cancer: identification and characterization using proteomic profiling and mass spectrometry. *Clin Cancer Res* (2003) 9(8):2904–11.
64. Kerbiriou D, Griffin J. Human high molecular weight kininogen. Studies of structure-function relationships and of proteolysis of the molecule occurring during contact activation of plasma. *J Biol Chem* (1979) 254(23):12020–7.
65. Golias C, Charalabopoulos A, Stagikas D, Charalabopoulos K, Batistatou A. The kinin system-bradykinin: biological effects and clinical implications. Multiple role of the kinin system-bradykinin. *Hippokratia* (2007) 11(3):124.
66. Montana V, Sontheimer H. Bradykinin promotes the chemotactic invasion of primary brain tumors. *J Neurosci* (2011) 31(13):4858–67. doi: 10.1523/JNEUROSCI.3825-10.2011
67. Ren Y, Ji N, Kang X, Wang R, Ma W, Hu Z, et al. Aberrant ceRNA-mediated regulation of KNG1 contributes to glioblastoma-induced angiogenesis. *Oncotarget* (2016) 5:1–12. doi: 10.18632/oncotarget.12659
68. Wang J, Wang X, Lin S, Chen C, Wang C, Ma Q, et al. Identification of kininogen-1 as a serum biomarker for the early detection of advanced colorectal adenoma and colorectal cancer. *PloS One* (2013) 8(7):e70519. doi: 10.1371/journal.pone.0070519
69. Yen C, Jiang S, Hsiao J, Chen C, Liu K. Identification of kininogen-1 as a potential prognostizc biomarker for oral cancer. *Eur J Cancer* (2016) 60:e10–e. doi: 10.1016/j.ejca.2016.03.039
70. Northrop-Clewes CA, Thurnham DI. Biomarkers for the differentiation of anemia and their clinical usefulness. *J Blood Med* (2013) 4:11. doi: 10.2147/JBM.S29212
71. Ludwig H, Pohl G, Osterborg A. Anemia in multiple myeloma. *Clin Adv Hematol Oncol: H&O* (2004) 2(4):233–41.
72. Sawayama H, Iwatsuki M, Kuroda D, Toihata T, Uchihara T, Koga Y, et al. Total iron-binding capacity is a novel prognostic marker after curative gastrectomy for gastric cancer. *Int J Clin Oncol* (2018) 23(4):671–80. doi: 10.1007/s10147-018-1274-7
73. Jacobs A, Slater A, Whittaker J, Canellos G, Wiernik P. Serum ferritin concentration in untreated Hodgkin's disease. *Br J Cancer* (1976) 34 (2):162–6. doi: 10.1038/bjc.1976.138
74. Vostrejs M, Moran PL, Seligman PA. Transferrin synthesis by small cell lung cancer cells acts as an autocrine regulator of cellular proliferation. *J Clin Investigation* (1988) 82(1):331–9. doi: 10.1172/JCI113591
75. Georgila K, Vyrla D, Drakos E. Apolipoprotein AI (ApoA-I), Immunity, Inflammation and Cancer. *Cancers* (2019) 11(8):1097. doi: 10.3390/cancers11081097
76. Lim LC, Looi ML, Zakaria SZS, Sagap I, Rose IM, Chin S-F, et al. Identification of differentially expressed proteins in the serum of colorectal cancer patients using 2D-DIGE proteomics analysis. *Pathol Oncol Res* (2016) 22(1):169–77. doi: 10.1007/s12253-015-9991-y
77. Fye HK, Wright-Drakesmith C, Kramer HB, Camey S, da Costa AN, Jeng A, et al. Protein profiling in hepatocellular carcinoma by label-free quantitative proteomics in two west African populations. *PloS One* (2013) 8(7):e68381. doi: 10.1371/journal.pone.0068381
78. Zabłocka-Słowińska K, Płaczkowska S, Skorska K, Prescha A, Pawełczyk K, Porębska I, et al. Oxidative stress in lung cancer patients is associated with altered serum markers of lipid metabolism. *PloS One* (2019) 14(4):e0215246. doi: 10.1371/journal.pone.0215246
79. Chung L, Moore K, Phillips L, Boyle FM, Marsh DJ, Baxter RC. Novel serum protein biomarker panel revealed by mass spectrometry and its prognostic value in breast cancer. *Breast Cancer Res* (2014) 16(3):R63. doi: 10.1186/bcr3676
80. Kozak KR, Su F, Whitelegge JP, Faull K, Reddy S, Farias-Eisner R. Characterization of serum biomarkers for detection of early stage ovarian cancer. *Proteomics* (2005) 5(17):4589–96. doi: 10.1002/pmic.200500093
81. Chen Y, Xiong X, Wang Y, Zhao J, Shi H, Zhang H, et al. Proteomic screening for serum biomarkers for cervical cancer and their clinical significance. *Med Sci Monitor: Int Med J Exp Clin Res* (2019) 25:288. doi: 10.12659/MSM.911478
82. Halton JM, Nazir DJ, McQueen MJ, Barr RD. Blood lipid profiles in children with acute lymphoblastic leukemia. *Cancer: Interdiscip Int J Am Cancer Soc* (1998) 83(2):379–84. doi: 10.1002/(SICI)1097-0142(19980715)83:2<379::AID-CNCR24>3.0.CO;2-P
83. Qin X, Chen Q, Sun C, Wang C, Peng Q, Xie L, et al. High-throughput screening of tumor metastatic-related differential glycoprotein in hepatocellular carcinoma by iTRAQ combines lectin-related techniques. *Med Oncol* (2013) 30(1):420. doi: 10.1007/s12032-012-0420-8
84. Goncalves A, Esterni B, Bertucci F, Sauvan R, Chabannon C, Cubizolles M, et al. Postoperative serum proteomic profiles may predict metastatic relapse in high-risk primary breast cancer patients receiving adjuvant chemotherapy. *Oncogene* (2006) 25(7):981–9. doi: 10.1038/sj.onc.1209131
85. Rižner TL. Discovery of biomarkers for endometrial cancer: current status and prospects. *Expert Rev Mol Diagn* (2016) 16(12):1315–36. doi: 10.1080/14737159.2016.1258302
86. Guo X, Hao Y, Kamilijiang M, Hasimu A, Yuan J, Wu G, et al. Potential predictive plasma biomarkers for cervical cancer by 2D-DIGE proteomics and Ingenuity Pathway Analysis. *Tumor Biol* (2015) 36(3):1711–20. doi: 10.1007/s13277-014-2772-5
87. Muntoni S, Atzori L, Mereu R, Satta G, Macis MD, Congia M, et al. Serum lipoproteins and cancer. *Nutrition Metab Cardiovasc Dis* (2009) 19(3):218–25. doi: 10.1016/j.numecd.2008.06.002
88. Chen Y-H, Magalhaes MC, et al. Hypoalbuminemia in patients with multiple myeloma. *Arch Internal Med* (1990) 150(3):605–10. doi: 10.1001/archinte.1990.00390150095018
89. Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* (2016) 44 (D1):D447–56. doi: 10.1093/nar/gkv1145

Conflict of Interest: Authors FR and DM are employed by Sciex, Gurgaon, India.

The remaining 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 Chanukuppa, Taware, Taunk, Chatterjee, Sharma, Somasundaram, Rashid, Malakar, Santra and Rapole. 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 Oncology

ORIGINAL RESEARCH
published: 18 February 2021
doi: 10.3389/fonc.2020.519180

Identifying Apoptosis-Related Transcriptomic Aberrations and Revealing Clinical Relevance as Diagnostic and Prognostic Biomarker in Hepatocellular Carcinoma

Jinyu Zhu[1,2†], Bufu Tang[1,2†], Xiuling Lv[3], Miaomiao Meng[3], Qiaoyou Weng[1,3], Nannan Zhang[1,2], Jie Li[1,2], Kai Fan[1,2], Liyun Zheng[1,3], Shiji Fang[1,3], Min Xu[1,3] and Jiansong Ji[1,3*]*

[1] Key Laboratory of Imaging Diagnosis and Minimally Invasive Intervention Research, Lishui Hospital, School of Medicine, Zhejiang University, Lishui, China, [2] Department of Radiology, Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China, [3] Department of Radiology, The Fifth Affiliated Hospital of Wenzhou Medical University, Lishui, China

In view of the unsatisfactory treatment outcome of liver cancer under current treatment, where the mortality rate is high and the survival rate is poor, in this study we aimed to use RNA sequencing data to explore potential molecular markers that can be more effective in predicting diagnosis and prognosis of hepatocellular carcinoma. RNA sequencing data and corresponding clinical information were obtained from multiple databases. After matching with the apoptotic genes from the Deathbase database, 14 differentially expressed human apoptosis genes were obtained. Using univariate and multivariate Cox regression analyses, two apoptosis genes (BAK1 and CSE1L) were determined to be closely associated with overall survival (OS) in HCC patients. And subsequently experiments also validated that knockdown of BAK1 and CSE1L significantly inhibited cell proliferation and promoted apoptosis in the HCC. Then the two genes were used to construct a prognostic signature and diagnostic models. The high-risk group showed lower OS time compared to low-risk group in the TCGA cohort ($P < 0.001$, HR = 2.11), GSE14520 cohort ($P = 0.003$, HR = 1.85), and ICGC cohort ($P < 0.001$, HR = 4). And the advanced HCC patients showed higher risk score and worse prognosis compared to early-stage HCC patients. Moreover, the prognostic signature was validated to be an independent prognostic factor. The diagnostic models accurately predicted HCC from normal tissues and dysplastic nodules in the training and validation cohort. These results indicated that the two apoptosis-related signature effectively predicted diagnosis and prognosis of HCC and may serve as a potential biomarker and therapeutic target for HCC.

Keywords: apoptosis, hepatocellular carcinoma, overall survival, relapse-free survival, diagnosis, prognosis

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Hamid Morjani, Université de Reims Champagne-Ardenne, France
Paul Tran, Augusta University, United States

***Correspondence:**
Jiansong Ji
lschrjjs@163.com
Min Xu
lschrxm@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *11 December 2019*
Accepted: *17 December 2020*
Published: *18 February 2021*

Citation:
Zhu J, Tang B, Lv X, Meng M, Weng Q, Zhang N, Li J, Fan K, Zheng L, Fang S, Xu M and Ji J (2021) Identifying Apoptosis-Related Transcriptomic Aberrations and Revealing Clinical Relevance as Diagnostic and Prognostic Biomarker in Hepatocellular Carcinoma. Front. Oncol. 10:519180. doi: 10.3389/fonc.2020.519180

INTRODUCTION

Primary liver cancer, 85–90% of which are hepatocellular carcinomas (HCCs), represents a major health burden being the sixth most commonly diagnosed cancer and the fourth leading cause of cancer death worldwide, and the incidence and mortality of HCC are dramatically increasing yearly (1–3). There are multiple management strategies used for treating HCC, each accompanied by different risks and rewards. For instance, the surgical options available are partial liver resection, percutaneous ablation, TACE, and either living or cadaveric donor transplantation, while radiation therapies are used for advanced HCC (4, 5). However, the relapse rates and adverse outcome rates of HCC remain high, and the 5-year relapse-free survival (RFS) and overall survival (OS) rates for HCC are still poor (6, 7). Therefore, reliable prognosis markers are strongly needed to establish dependable prognosis models and make accurate survival predictions, which may help to specify individualized treatment options. The molecular pathogenesis of HCC is very complex, with multiple genetic variations, such as mutations and deletions of genes, and alterations in key molecular regulatory pathways (8). Based on this status quo, it is very urgent to explore new molecular markers from the molecular mechanism level to improve the accuracy of prognosis and survival prediction of HCC.

Apoptosis, as an important and tightly governed process of programmed cell death (PCD) with distinct biochemical and genetic pathways, involving the genetically determined elimination of cells (9). Apoptosis not only occurs as a fundamental homeostatic mechanism in the maintenance of normal cell turnover in tissues during development and aging (10), but previous research has revealed that dysfunctional apoptosis (too little or too much apoptosis or apoptosis occurring in the wrong place/time) also leads to variable diseases and multiple types of cancer (11). Many human diseases can be directly or indirectly attributed to the disorder of the apoptotic mechanism, which, triggered by abnormalities in the apoptosis pathway, lead to the excessive accumulation of cells or the loss of the disease (12). Previous studies have reported that the apoptotic pathway is important in controlling the critical processes of HCC. For instance, Changqing Su et al. (13) revealed that survivin, as a member of the inhibitor of apoptosis protein (IAP) family, is highly expressed in HCC, and its overexpression can promote cancer cell proliferation, inhibit cancer cell apoptosis, reduce the sensitivity of cancer cells to radiotherapy and chemotherapy, and ultimately affect the prognosis of patients with HCC. Few studies have explored the association between apoptotic genes and the diagnosis and prognosis of HCC through high-throughput sequencing of biomarkers.

In this study, we aimed to explore key apoptosis-related genes that can help predict the diagnosis and prognosis of HCC more accurately as new molecular markers. We explored the differences in mRNA expression between HCC tissues and adjacent noncancerous tissues from the TCGA database to determine the apoptotic-related genes involved in the prognosis of liver cancer. Functional enrichment analysis of these apoptosis-related genes was performed. Through univariate and multivariate Cox regression analysis, we eventually identified and constructed a prognostic model of two apoptosis-related genes, which contained BAK1 and CSE1L. Knockdown of BAK1 and CSE1L were showed to inhibit the proliferation of HCC cells, and significantly promote the apoptosis of HCC. And two diagnostic models were constructed by the same genes to distinguish HCC from normal samples and dysplastic nodules. Moreover, a predictive nomogram based on survival was constructed and evaluated to improve the prognostic and survival prediction accuracy of HCC patients and help clinicians make more useful clinical decisions. We also validated the expression patterns of the two genes in the Gene Expression Omnibus (GEO), TCGA cohort, ICGC cohort, and online databases including UALCAN and Kaplan-Meier Plotte. GSEA analysis was performed to explore the underlying mechanism of the two genes in HCC.

MATERIALS AND METHODS

Apoptosis-Related Genes

Fifty-four apoptosis genes were extracted from the Deathbase (http://deathbase.org), which contains a series of proteins and corresponding coding genes of various species involved in cell death.

Selection of HCC Associated RNA-Sequencing Data

The RNA sequence data (Illumina HiSeq RNA-Seq platform) and corresponding clinical information of liver hepatocellular carcinoma (LIHC) patients were obtained from the Cancer Genome Atlas (TCGA) data portal. The LIHC cohort contained 370 HCC samples and 50 normal tissue samples as of July 8, 2019. Considering that the RNA sequence data were downloaded from TCGA, they are open to the public and are freely available, and this study strictly followed the access policy and publishing guidelines of the TCGA database, so there was no need to seek approval from the Ethics Committee for this research.

Differentially Expressed mRNA Screening Between HCC and Adjacent Non-Cancerous Tissues

First, we matched the 54 apoptotic genes obtained from the Deathbase database with 19,677 annotated coding genes of HCC downloaded from the RNA-sequencing platform in the TCGA database, from which we obtained 45 human apoptosis genes associated with HCC. Then, for subsequent analysis, we used the limma. R package to identify 14 differentially expressed genes (DEGs) of the dataset with absolute log2 fold change (FC) >1 and adjusted P value <0.05.

Functional Enrichment Analysis

The Database for Annotation, Visualization and Integration Discovery (DAVID) (https://david.ncifcrf.gov) was used to analyze The Gene Ontology (GO), and KOBAS database

(kobas.cbi.pku.edu.cn) was used to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of differentially expressed HCC apoptosis-related genes. Adjusted $P < 0.05$ was considered to represent statistical significance.

Construction of the Apoptosis-Related Prognostic Model

Next, we validated the correlation between OS of HCC patients and specific expression levels of each apoptosis-related gene by univariate and multivariate Cox regression analysis. In the univariate Cox regression analysis, a P value for a gene <0.05 was considered to represent statistical significance, and the gene was also considered to be significantly associated with the OS of HCC patients. Subsequent multivariate Cox regression analysis further screened two apoptosis-related genes that could be deemed crucial prognostic factors for evaluating survival, thereby ultimately constructing the optimal apoptosis-related prognostic signature. Prognostic risk scores based on two apoptosis-related genes were established by regression coefficients of the multivariate Cox regression model (β) and gene expression levels. Prognostic index (PI) = (β* expression level of BAK1) + (β* expression level of CSE1L). Setting the median of the PI as a cut-off value, 370 HCC patients were divided into a high-risk group and a low-risk group from the prognostic model.

Build of the Apoptosis-Related Diagnostic Model

Logistic regression analysis was applied in building a diagnostic model to distinguish HCC from normal samples and dysplastic nodules. The P value is defined as the probability that patient may have HCC. In order to make the P value between 0 and 1, we performed logistic transformation on P and use logit(P) as the dependent variable, that is, $\text{Logit}(P) = \ln[p/(1-p)] = \alpha+\beta_1 x_1+\beta_2 x_2$. x_1 represents the expression level of BAK1 and x_2 represents the expression level of CSE1L. Based on the expression data of BAK1 and CSE1L in HCC patients and normal subjects from the TCGA database, the parameters are estimated, and the coefficients of the logistic regression equation are estimated using SPSS (version 24.0) to obtain the logit(P) equation. If P of a subject calculated based on the expression of BAK1 and CSE1L is between 0 and 0.5, it is classified as a healthy person, otherwise it is classified as a patient with HCC.

Establishment and Validation of a Predictive Nomogram

To verify whether the predictive power of this apoptosis-related prognostic signature was independent of other clinical features of HCC patients, which included age, sex, weight, AFP, vascular tumor cell invasion, clinical stage, and pathological grade, we performed univariate and multivariate Cox regression analysis, setting the prognostic signature and clinical features as independent variables, and OS as the dependent variable. Bilateral P values <0.05 were considered to represent statistical significance. Then, the hazard ratio and 95% confidence interval (CI) were calculated for each clinical variable.

Based on these independent prognostic factors, a predictive nomogram (including important independent prognostic factors and calibration maps) was constructed using rms R software. The nomogram was then evaluated and validated by calibration and discrimination. In addition, a consistency index (C index) is calculated using a bootstrap method with 1,000 resamples to determine the distinction of the nomogram. Meanwhile, we used the C index, time-dependent receiver operating characteristic (ROC) analysis and decision curve analysis (DCA) to compare the prediction accuracy between this nomogram and other nomograms constructed by single individual prognostic factors. DCA is a novel tool for evaluating predictive models that quantifies the clinical applicability of nomograms by analyzing the clinical outcomes resulting from decisions made based on predictive nomograms and is of great value in determining alternative diagnostic and prognostic strategies (14). P value <0.05 was considered to represent statistical significance.

Cell Culture and siRNA Treatment

The human hepatocellular carcinoma cell line and human L02 hepatocytes were obtained from the American Type Culture Collection (ATCC) (Manassas, VA, USA). Cells cultured in DMEM media supplemented with 2 mM L-glutamine, 100 U/ml (penicillin-streptomycin), and 10% fetal bovine serum (FBS) in 5% CO2 at 37°C in a constant temperature incubator. SK-HEP1 and SMMC-7721 was transfected 50 nM BAK1 and CSE1L siRNA in basic DMEM media for 6–8 h before change with DMEM media supplemented FBS and penicillin-streptomycin.

Western Blot Analysis

Protein was collected from SK-HEP1 and SMMC-7721 cells using protein lysis buffer containing PMSF (Bio-Rad). The BCA method was used to determine the total protein concentration. After adding the loading buffer to the protein samples, heating the protein samples at 100°C for 5 min. Protein samples were separated using 10% sodium lauryl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then transferred to a PVDF membrane. Then 5% skim milk was formulated to block the blot for 1.5 h at room temperature on a shaker. The blots were incubated with specific primary antibodies on a shaker overnight at 4°C. Imprinted with horseradish peroxidase-conjugated secondary antibody and processed with ECL kit (Bio-Rad) to detect protein.

RNA Preparation, Reverse Transcription, and qRT-PCR

Total RNA was extracted using the TRIzol® reagent (Invitrogen, shanghai, China), and the cDNA was prepared by TransScript Top Green qPCR supermix (TransGen, Guangzhou, China) according to the manufacturer's protocol. Relative gene expression was performed with RT-PCR detection system (Bio-Rad Laboratories, Inc., Hercules, CA, USA) with Real-Time PCR Master Mixes (TransGen, Guangzhou, China) according to the manufacturer's protocol. β-actin was served as an endogenous control. The relative expression (fold change) of target mRNA was measured using the 2-ΔΔCT methodology. Each experiment was repeated three times independently.

Cell Proliferation Assay

After transfecting BAK1 and CSE1L siRNA for 48 h, SK-HEP1 and SMMC-7721 were cultured in 96-well plates for 48 h, the cell number was 3,000/plate with 200 ul DMEM. Then CCK8 agent (Yeasen, Shanghai, China) was treated per plate and OD450 was analyzed by Micro plate spectrophotometer (Thermo Fisher Scientific, MA, USA). Edu assay was also performed to measure the cell proliferation. Edu agent was obtained from Ruibo company (Ruibo, Guangzhou, China) and the Edu assay is performed according to the manufacturer's protocol.

Flow Cytometry to Measure Apoptotic Cells

After transfecting BAK1 and CSE1L siRNA for 72 h, Flow cytometry was used to detect the apoptosis of SK-HEP1 and SMMC-7721 cells. Annexin and PI agent (Yeasen, Shanghai, China) was used to label the viable apoptotic cells and non-viable apoptotic cells according to the manufacturer's protocol. Flow cytometry was used to measure the number of apoptotic cells. All samples were analyzed using NovoExpress installed with software.

Statistical Analysis

Statistical analyses were performed using SPSS, version 22.0 (SPSS, Chicago, IL, USA). Values were expressed as the mean ± SD. Quantitative data in paired groups were determined using the Student's t-test. One-way analysis of variance was performed for multiple group comparisons. A value of $P < 0.05$ indicated significant differences.

RESULTS

Identification and Validation of the Key Apoptotic Genes Closely Related to OS in HCC Patients

By comparing the differentially expressed genes in mRNA expression profiles between HCC tissues (n = 370) and adjacent non-cancerous tissues (n = 50) and matching with apoptotic-related genes (**Table S1**), 14 differentially expressed human apoptotic genes (log FC>1 or log FC<-1, adjusted $P < 0.05$) adjusted for false discovery rate were obtained (**Figures 1A, B**). All the genes were overexpressed in the HCC tissues (**Figure 1C**). We then further

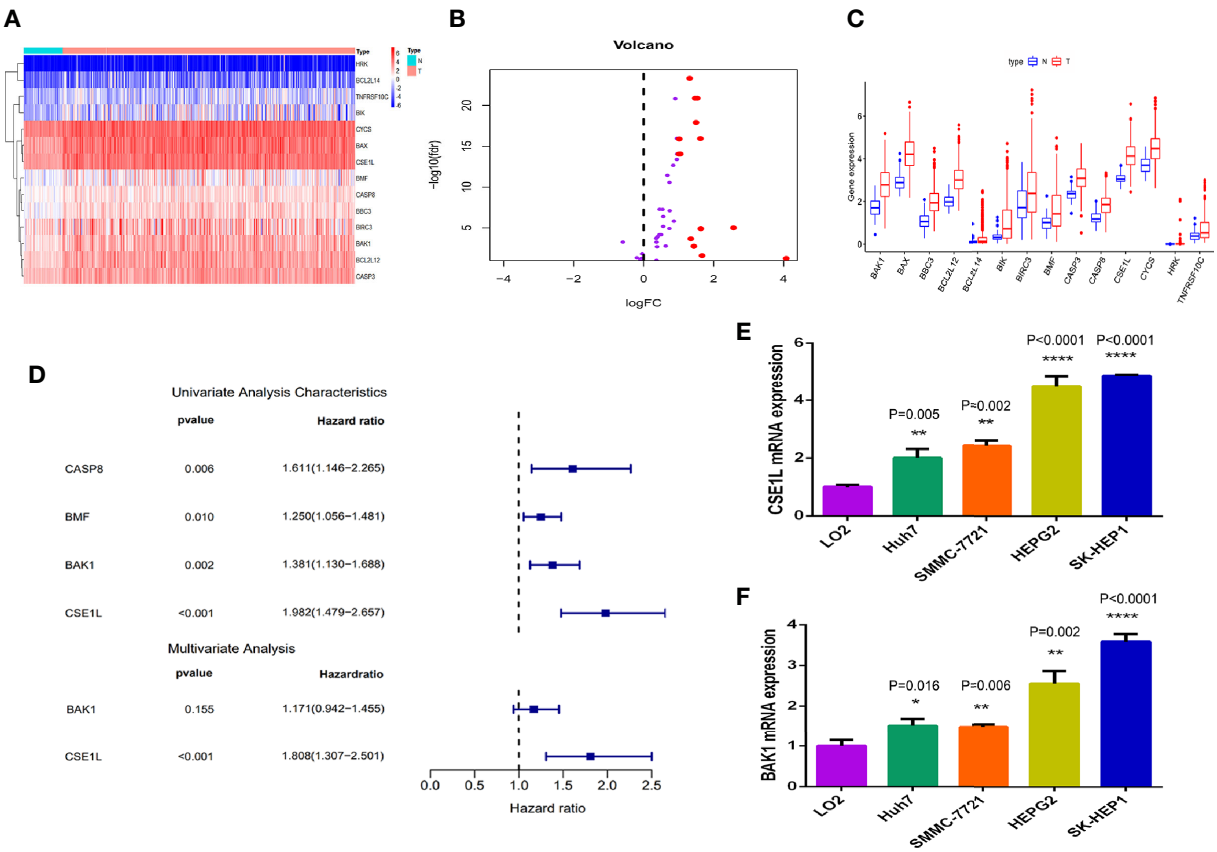


FIGURE 1 | Differential expression analysis and identification of key apoptosis genes related to prognosis in HCC. **(A)** The heat map showed the expression levels of 14 apoptotic genes in HCC tissues and noncancerous tissues. **(B)** The histogram indicated the expression patterns of 14 apoptotic genes in HCC tissues and noncancerous tissues. **(C)** The volcano plot indicates that the 14 differentially expressed genes satisfy the condition of $P < 0.05$ and log FC > 1. **(D)** Univariate and multivariate Cox regression analysis performed to identify the genes closely associated with the prognosis of HCC patients. **(E, F)** qRCR was used to measure the mRNA level of CSE1L **(E)** and BAK1 **(F)** in human L02 hepatocytes and HCC cell lines. *P < 0.05; **P < 0.01; ****P < 0.0001.

explored the relevance of these differentially expressed apoptotic genes to the prognosis of HCC.

Next, we investigated the potential biological processes and pathways of these 14 differentially expressed apoptotic genes. We performed a GO function enrichment analysis of these genes through the DAVID database and finally identified 10 important GO terms (adjusted $p < 0.05$). The association between apoptotic genes and GO terms was visualized through the GOplot.R package (**Figure S1A**). According to the degree of gene function enrichment, the top five GO terms were "positive regulation of release of cytochrome c from mitochondria," "positive regulation of apoptotic signaling pathway," "regulation of apoptotic signaling pathway," "positive regulation of mitochondrion organization," and "regulation of release of cytochrome c from mitochondria." Subsequently, we performed a KEGG pathway analysis concentrating on these genes through the KOBAS database and finally determined 26 statistically significant KEGG pathways (adjusted $P < 0.05$). We then visualized the association between apoptotic genes and KEGG through the enrichplot.R package (**Figure S1B**), and the top three important pathways among them were "apoptosis-multiple species," "apoptosis," and "platinum drug resistance."

To analyze the association between these 14 differentially expressed apoptotic genes and OS in HCC patients, we performed a univariate Cox regression analysis and screened four apoptotic genes (CASP8, BMF, BAK1, CSE1L) closely associated with OS in HCC patients ($P < 0.05$) (**Figure 1D**). Then, through multivariate Cox regression analysis, two apoptotic genes (BAK1 and CSE1L) were identified to be significantly associated with OS in HCC patients. To explore the role of two apoptotic genes in HCC, qPCR assay was used to measure the mRNA of BAK1 and CSE1L in different cell lines. We found BAK1 and CSE1L level in HCC cells was obviously higher than in human L02 hepatocytes (**Figures 1E, F**).

Identified the Tumoral Effect of Critical Apoptotic Regulators in Hepatocellular Carcinoma Cells

To further investigate explore the oncogene role of the two critical apoptotic regulators in Hepatocellular Carcinoma cells. Then we knocked down the level of BAK1 and CSE1L in HCC cells including SMC-7721 and SK-HEP1. Western blot showed the expression of BAK1 and CSE1L was obviously inhibited with BAK1 and CSE1L siRNA administration (**Figures 2A–D**). The CCK assay showed BAK1 inhibition suppressed HCC cells proliferation (**Figures 2E, G**) and the anti-proliferation effect of CSE1L inhibition was similar to BAK1 knockdown in HCC cells (**Figures 2F, H**). In addition, Edu assay indicated BAK1 and CSE1L inhibition significantly suppressed HCC cells proliferation in SMC-7721 and SK-HEP1 cell (**Figures 2I, J**). Moreover, flow cytometry showed BAK1 and CSE1L suppression promoted cell apoptosis in SMC-7721 and SK-HEP1 cells (**Figures 2K, L**), and quantitative statistical results are presented in **Figures 2M–P**. The above results confirmed that BAK1 and CSE1L inhibited HCC cells apoptosis and were closely related to the HCC cells proliferation.

Construction of the Apoptosis-Related Prognostic Signature and Evaluation of Its Predictive Performance of Prognosis in HCC Patients in TCGA Cohort

An apoptosis-related prognostic signature based on BAK1 and CSE1L was established. The predictive index (PI) = (0.157511873* the expression level of BAK1) + (0.592406034* the expression level of CSE1L). The median PI was set to the cut-off value, and all 370 HCC patients were divided into a high-risk group (n = 185) and a low-risk group (n = 185) (**Figure S2A**). The corresponding relationship between the patient's risk score and survival status is shown in **Figure 3A**, revealing that patients with low-risk scores had better survival rates than those with high-risk scores ($P < 0.05$). The expression levels of BAK1 and CSE1L were different in HCC patients with different risk scores (**Figure S2B**). Using the HCC patients in the TCGA cohort, the K-M curve was used for survival analysis. The high-risk group showed lower survival time compared to low-risk group ($P < 0.001$, HR = 2.11) and the OS probabilities at 1 year, 3 years, and 5 years were 0.72, 0.48, and 0.37 in high-risk patients and 0.926, 0.73, 0.56 in low-risk patients (**Figure 3B**). The recurrence probabilities of high-risk patients were higher than those of low-risk patients ($P = 0.003$) (**Figure 3D**). The AUCs of the prognostic signature were 0.7, 0.76, 0.65, and 0.62 at 0.5 year, 1 year, 3 years, and 5 years, respectively (**Figure 3C**), suggesting that the prognostic signature has good predictive power.

External Validation of the Predictive Power of This Apoptosis-Related Prognostic Signature for Prognosis in HCC Patients in the GEO Cohort and ICGC Cohort

To further confirm that the apoptosis-related prognostic signature has good predictive value in terms of prognosis in HCC patients, the prognostic signature was applied to the GSE14520 cohort (containing 243 tumor samples and 241 adjacent non-tumor samples) and ICGC cohort (including 243 tumor samples and 202 adjacent non-tumor samples). Using the same cut-off value obtained in the TCGA cohort, the GSE14520 cohort and ICGC cohort were divided into a high-risk group and a low-risk group (**Figures S2C, E**). **Figures 3E, I** show that patients with low-risk scores in this data set had better survival probabilities than those with high-risk scores, and these results are consistent with the TCGA cohort. The expression levels of BAK1 and CSE1L in patients with different risk scores are shown in **Figures S2D, F**. K-M curve analysis was performed on HCC patients in the GSE14520 cohort and ICGC cohort, the high-risk group showed lower survival time compared to low-risk group in GSE14520 cohort ($P = 0.003$, HR = 1.85) and ICGC cohort ($P < 0.001$, HR = 4) (**Figures 3F, J**). Moreover, high-risk patients showed higher recurrence rates in GSE14520 cohort (**Figure 3H**). The AUCs in the time-dependent ROC curve analysis reached 0.7, 0.66, 0.58, and 0.56 of GSE14520 cohort (**Figure 3G**) and 0.65, 0.72, 0.7, and 0.79 at of ICGC cohort (**Figure 3K**) at 0.5 year, 1 year, 3 years, and 4 years, suggesting that the prognostic signature has good predictive sensitivity and

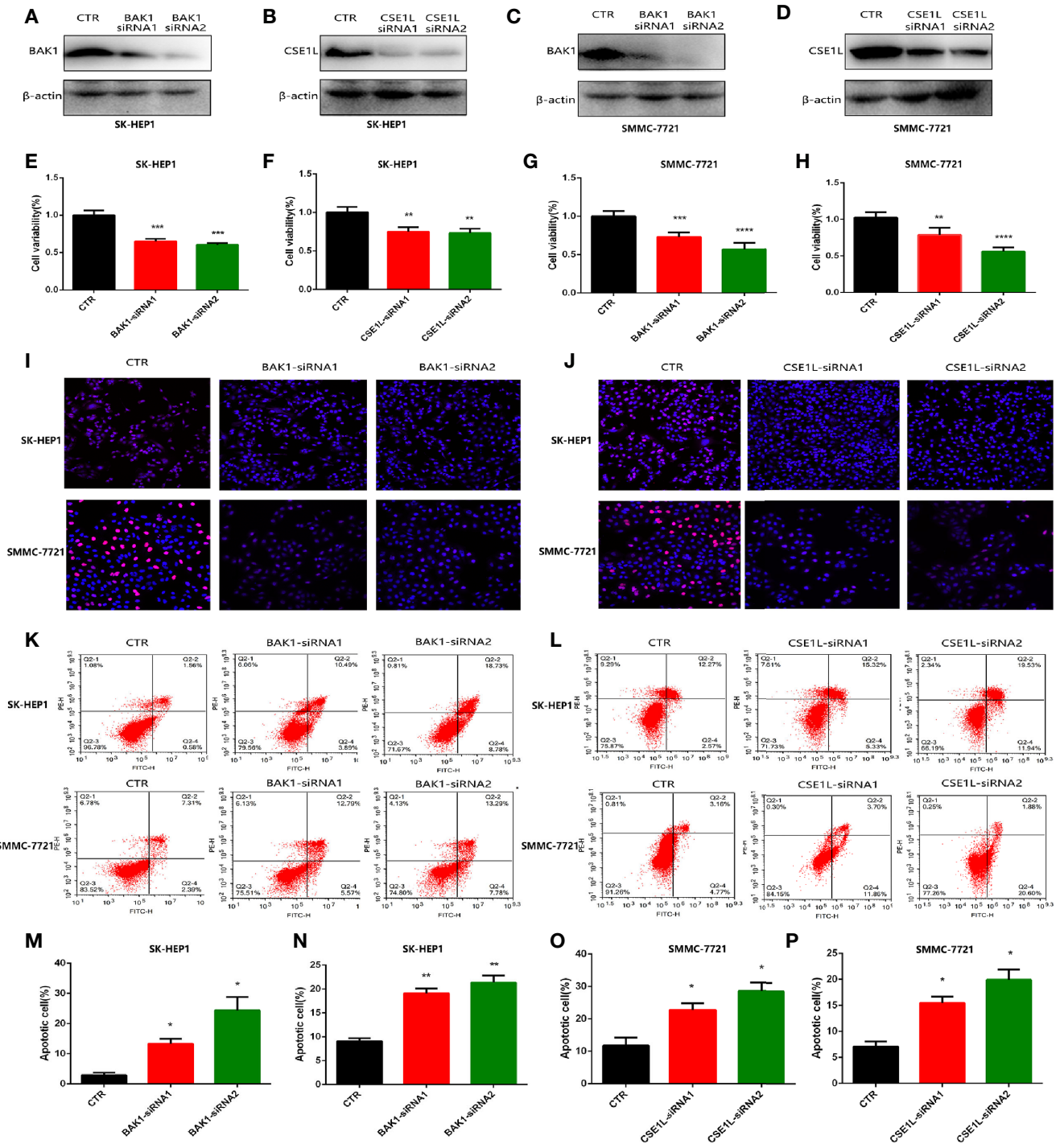


FIGURE 2 | Measure the oncogenic effect of BAK1 and CSE1L in the HCC. **(A–D)** Western blot determined that expression of BAK1 and CSE1L was inhibited with BAK1 and CSE1L siRNA administration. **(E–H)** CCK8 assay showed BAK1 and CSE1L inhibition significantly suppressed HCC cells proliferation. **(I, J)** Edu assay indicated knock down BAK1 and CSE1L inhibited HCC cells proliferation, respectively. **(K, L)** Flow cytometry confirmed decreased BAK1 and CSE1L expression promoted HCC cells apoptosis in the HCC. **(M–P)** Quantitative statistical results of the effects of BAK1 and CSE1L expression on cell apoptosis. Data are shown as the mean ± SD of at least three independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

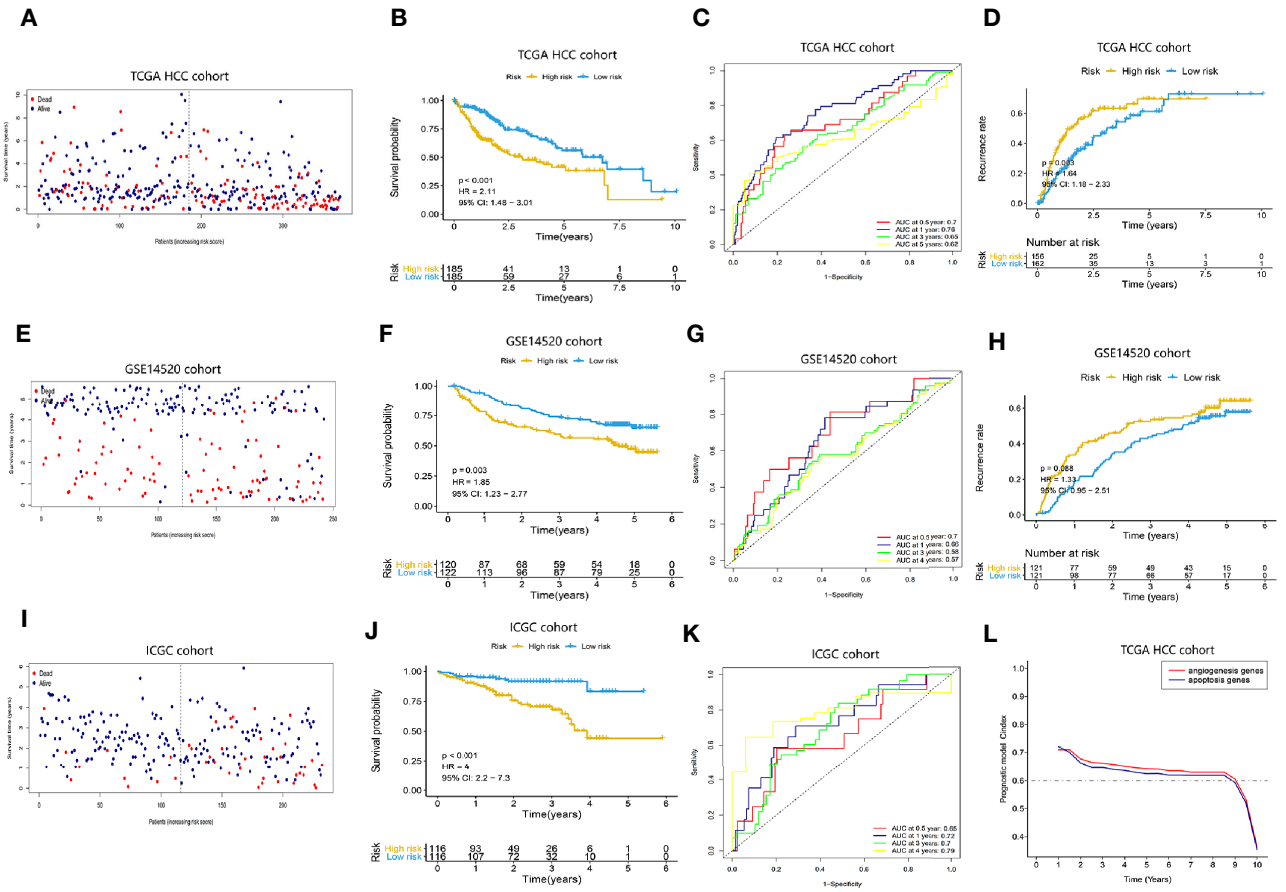


FIGURE 3 | Assessment of the predictive performance of the apoptosis-related prognostic signature. Kaplan-Meier survival analysis, heat map and time-dependent ROC curves of the prognostic signature in the TCGA HCC cohort, GEO HCC cohort, and ICGC cohort. **(A, E, I)** The survival status of the HCC patients with different risk scores. **(B, F, J)** K-M curves indicate that OS in high-risk patients is significantly lower than in low-risk patients (P < 0.05). **(C, G, K)** Time-dependent ROC curve analysis of the apoptosis-related prognostic signature. **(D, H)** K-M curves showed the difference of recurrence rates between high-risk patients and low-risk patients. **(L)** The difference in predictive performance of the two prognostic models.

specificity. The result suggested the prognostic model can be applied to different platforms and has good clinical value in GSE14520 cohort and ICGC cohort. And compared with our previous research (15), we confirmed that the apoptosis-related prognostic model performed better prediction accuracy (**Figure 3L**)

Hierarchical Analysis of Predictive Performance of OS in This Apoptosis-Related Prognostic Signature Based on Clinical Stage

Consistent with the apoptosis-related prognosis signature, the TNM stage was also significantly correlated with the prognosis and survival status of HCC patients. The survival status resulting from the progression of HCC indicated that the risk index of the dead patients was significantly higher than that of the alive patients (P = 0.004) (**Figure 4A**). And the prognostic risk index of patients in TNM stages III and IV was determined to be significantly higher than that of patients in stages I and II (P = 0.005) (**Figure 4B**). Hierarchical analysis was used to determine whether the prognostic performance of the prognostic signature was independent of clinical TNM stage. HCC patients in TCGA cohort were divided into two groups according to TNM stage: I, II, and III, IV. The hierarchical analysis authenticated that the prognostic signature was significantly associated with OS in the HCC cohorts of both groups (**Figures 4C, D**), indicating that the prognostic signature has good predictive performance for OS in HCC patients, which was independent of clinical TNM stage.

The Apoptosis-Related Prognostic Signature Is Independent of Traditional Clinical Features in Predicting OS

In order to assess the independent predictive value of this apoptosis-related prognostic signature, univariate and multivariate Cox regression analyses were used to analyze 370 HCC patients with intact clinical features from the TCGA cohort. Univariate Cox regression analysis revealed that age,

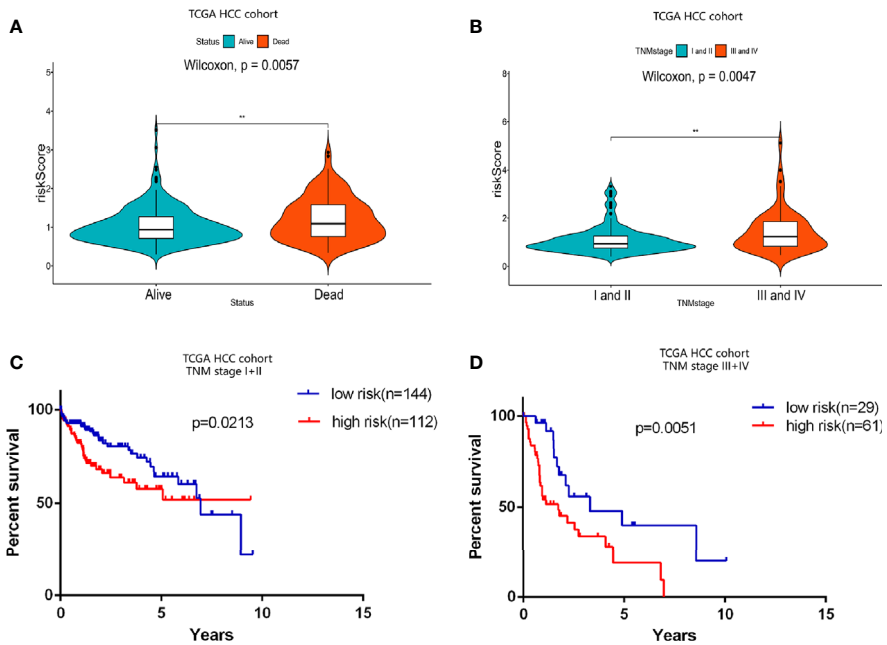


FIGURE 4 | Prognostic analysis of the apoptosis-related prognostic signature according to clinical TNM stage. **(A)** Violin plot revealing significantly different risk scores between patients in stages I and II and patients in stages III and IV (**P < 0.01). **(B)** Violin plot revealing significantly different risk scores between alive patients and dead patients (**P < 0.01). **(C, D)** Kaplan-Meier (K-M) survival analysis of different clinical TNM stages.

clinical stage, and the apoptosis-related prognostic signature were independent prognostic factors associated with OS in HCC patients (P < 0.05) (**Figure 5A**) (**Table S2**).

For more accurate prediction of the prognosis and survival probability of HCC patients and to help clinicians make clinical decisions and develop personalized treatment plans for patients, we constructed a predictive nomogram (**Figure 5B**). Compared with age and clinical stage, the risk score of the prognostic signature had the greatest impact on total points, which is consistent with the results of multivariate Cox regression analysis. The C index of the nomogram was 0.6777 after 1,000 bootstrap cycles (95% CI 0.61, 0.75), demonstrating that the prediction results of the nomogram were consistent with the actual results. The 1-year, 3-year, and 5-year calibration curves in the calibration chart were found to be very close to the best predictive curve (45° line), indicating that the predictive performance of the nomogram on prognosis was good (**Figures 5C–E**).

In addition, we compared the predictive accuracy between the nomogram and a single independent prognostic factor (age, clinical stage, or the apoptosis-related prognostic signature). We found that the C index of the nomogram (0.68) was superior to the C index of age (0.57), clinical stage (0.54), or prognostic model (0.66), confirming that the nomogram had better predictive value than any single independent prognostic factor. The ROC curve showed that the AUCs of the nomogram at 1 year, 3 years, and 5 years were all the largest (**Figures 6A–C**), demonstrating that the nomogram had the best predictive performance. The results of DCA also revealed that the nomogram provided the best net benefit compared to any single independent prognostic factor (**Figures 6D–F**). These findings demonstrate that the nomogram was more accurate than any single independent prognostic factor in predicting OS in HCC patients.

Construction and Validation of the Diagnostic Model Based on Two Apoptosis Genes

To confirm the diagnostic value of the two apoptosis-related genes, we established a diagnostic model to distinguish HCC and normal samples. The diagnostic score = −8.525 + (0.317* the expression level of BAK1) + (0.923* the expression level of CSE1L). According to the diagnostic model, the predicted specificity and sensitivity was 0.92 and 0.88 in the TCGA cohort (including 50 paired HCC and normal samples) (**Figure 7A**), and the specificity and sensitivity in the ICGC cohort (containing 243 HCC and 202 normal samples) was 0.80 and 0.89 (**Figure 7B**). ROC curve analysis showed the AUC was 0.956 in the TCGA cohort (**Figure 7C**) and 0.913 in the ICGC cohort (**Figure 7D**), respectively. Unsupervised hierarchical clustering indicated the two gene model separated HCC from normal samples with higher specificity and sensitivity (**Figures 7E, F**).

Dysplastic nodules less than 2 cm in diameter were not well detected by CT or MRI. To further evaluate the diagnostic value of the two genes, we constructed diagnostic model to distinguish HCC from dysplastic nodules. The diagnostic score = −63.824 + (2.03* the expression level of BAK1) + (9.21* the expression level of CSE1L). The diagnostic model showed specificity at 0.76 and

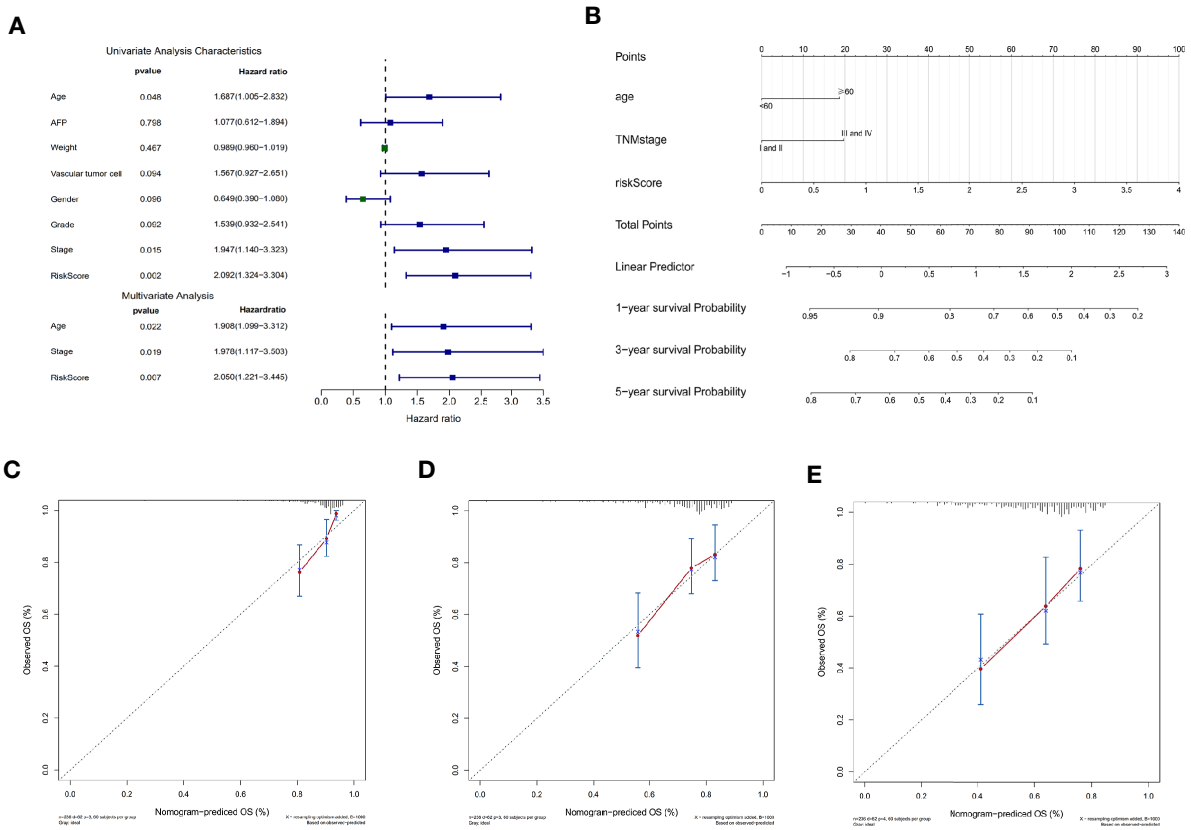


FIGURE 5 | Cox regression analysis and nomogram incorporating three independent prognostic factors was constructed to predict OS in HCC patients. **(A)** Univariate and multivariate Cox regression analysis of the apoptosis-related prognostic signature and clinical features with overall survival (OS) in HCC patients ($P < 0.05$). **(B)** The total points were obtained by cumulatively adding the corresponding points of the independent prognostic factors on the point scale to predict the OS of the HCC patients at 1 year, 3 years, and 5 years in the nomogram. **(C–E)** The calibration curve to assess the predictive value of the nomogram in OS at 1 year, 3 years, and 5 years. The Y-axis represents the actual OS measured by K-M curve analysis, and the X-axis represents the OS predicted by the nomogram.

sensitivity at 0.94 in training cohort (GSE 6764) (**Figure 8A**), and specificity at 0.92 and sensitivity at 0.84 in validation (GSE 98620) (**Figure 8B**). The AUC of the diagnostic models were 0.933 and 0.812 in the two cohort (**Figures 8C, D**). Unsupervised hierarchical clustering showed the two gene models had good diagnostic value of distinguishing HCC from dysplastic nodules (**Figures 8E, F**). These data suggested two apoptotic genes were potential diagnostic markers for HCC.

Further Validation of the Expression of Two Apoptosis Genes and Their Predictive Power for OS and RFS

To further validate the value of the two apoptosis-related genes (BAK1 and CSE1L) for constructing the prognostic and diagnostic signatures, we adopted HCC cohorts from ICGC database and the GEO database (GSE14520) in to analyze the expression levels of these two apoptotic genes. The results of validation analysis in GSE14520 (**Figures 9A, B**) and ICGC (**Figures 9C, D**) are consistent with the finding in UALCAN database. Survival analysis in Kaplan-Meier Plotte (16) revealed that the OS of HCC patients with high expression of BAK1 was significantly lower than that of patients with low expression of BAK1 ($P = 0.001$) in TCGA cohort (**Figure 9E**) while there was no significant difference in RFS ($P = 0.13$) (**Figure 9F**). HCC patients with high CSE1L expression had significantly lower OS and RFS than patients with low expression of CSE1L ($P < 0.05$) (**Figures 9G, H**). K-M curves we performed in HCC cohort from TCGA is in accordance with the analysis result in Kaplan-Meier Plotte (**Figures 9I–L**). In the HCC cohort from GSE14520 and TCGA, The AUCs in time-dependent ROC curve analysis confirmed that BAK1 and CSE1L have great predictive value of prognosis (**Figures 9M–P**).

Exploration of the Potential Mechanism of BAK1 and CSE1L in the HCC

For exploring underlying mechanism of two apoptosis genes in HCC, correlation analysis was performed and indicated BAK1 expression had positive correlation with CSE1L in HCC in TCGA cohort and ICGC cohort (**Figures 10A, B**). PPI network showed the potential regulator which interacted with BAK1 and CSE1L (**Figures 10C, D**). In addition, GSEA analysis showed BAK1 and CSE1L regulated underlying signal pathway

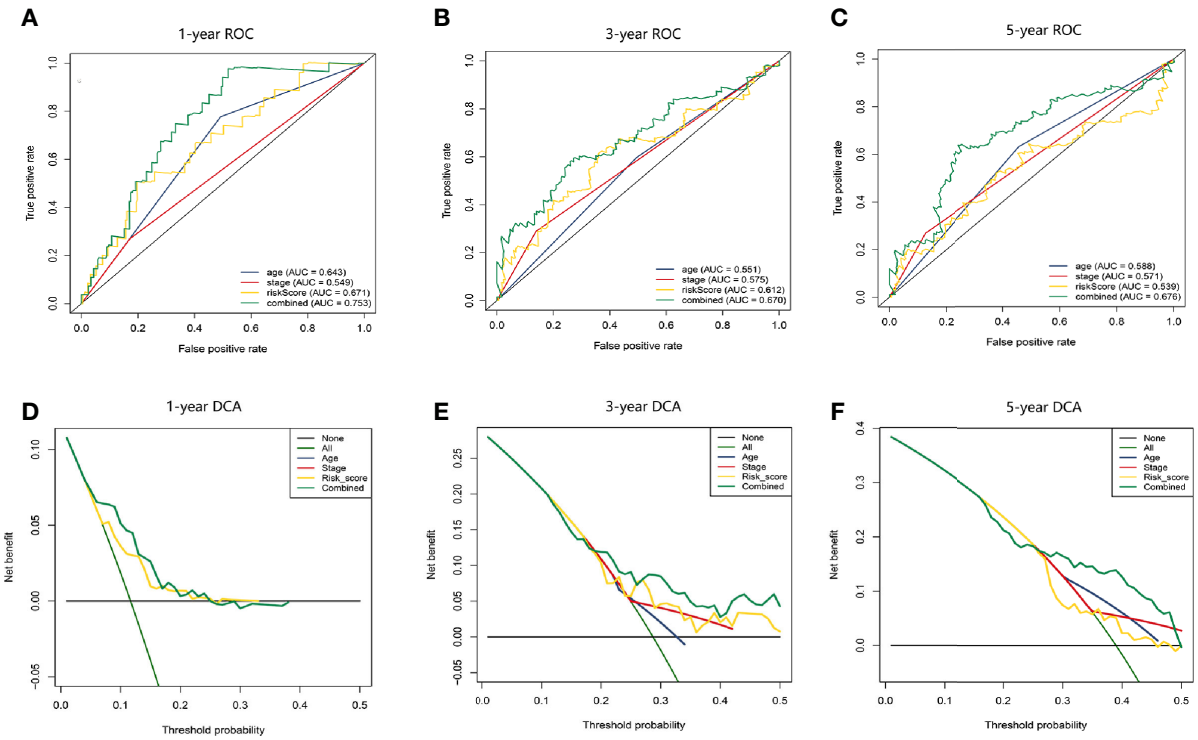


FIGURE 6 | The time-dependent ROC and DCA curves were drawn to compare the predictive accuracy for prognosis between different nomograms. **(A–C)** Time-dependent ROC curve analysis validated the predictive accuracy of the nomograms constructed from different independent prognostic factors at 1 year, 3 years, and 5 years. **(D–F)** DCA visually assesses the clinical benefits of nomograms and the range of applications in which the nomogram can yield clinical benefit at 1 year, 3 years, and 5 years.

in HCC (**Tables S3**, **S4**). According to the top 10 signal pathway based on GSEA analysis, we found BAK1 and CSE1L also played positively role in signaling pathway which is relatively correlation with HCC including pathway in cancer. Moreover, BAK1 and CSE1L played a common role in all the top 10 negatively correlated pathways (**Figure 10E**). The results suggested BAK1 and CSE1L may played a synergistic role in the progression and development of hepatocellular carcinoma.

DISCUSSION

Due to the complex molecular pathogenesis, HCC is one of the most deadly malignant tumors in the world, with high mortality and recurrence rates and poor disease outcomes (1, 17). The recurrence of HCC is closely related to the early diagnosis and prognosis, but there is no consistent early diagnosis and prognostic examination standard so far. The diagnosis of HCC is rely on pathological (18) or radiological evaluation (19, 20), and the prognosis of HCC is mainly evaluated by AJCC TNM classification and BCLC staging currently (21). However, these methods of assessing the diagnosis and prognosis of HCC patients are far from satisfactory. Some studies have shown that a combination of molecular markers may have good value in evaluating the diagnosis and prognosis of cancer patients, which may be more helpful in predicting clinical outcomes (22–24). Realizing that biomolecular markers may provide a more effective predictive tool for tumor diagnosis and prognosis in clinical practice, comprehensive genomic research at the RNA level has received extensive attention. However, the number of molecular markers that have been identified as diagnosis and prognostic factors are still very small. In order to reduce the mortality and recurrence rates, improve clinical outcomes and survival, and help clinicians more effectively specify individualized treatment strategies, we urgently need to screen out new molecular markers that can predict diagnosis and prognosis more accurately.

Apoptosis, also identified as programmed cell death, is an important component of the physiological process to regulate the renewal or death of normal cells (9). However, dysregulation of apoptosis plays a critical role in the occurrence of many diseases and even multiple cancers. During the development and progression of cancer, alterations in the intracellular apoptotic pathway induce a reduction in apoptosis, and decreased apoptosis prevents malignant cells from dying, which promotes the development of cancer and is associated with resistance to treatment (25, 26). Correspondingly, apoptosis can also play an important role in cancer treatment as a therapeutic target, affecting the prognosis of tumors. Most of the current research has focused on the role of single apoptosis-related genes in the diagnosis and prognosis of HCC. Few studies have explored the

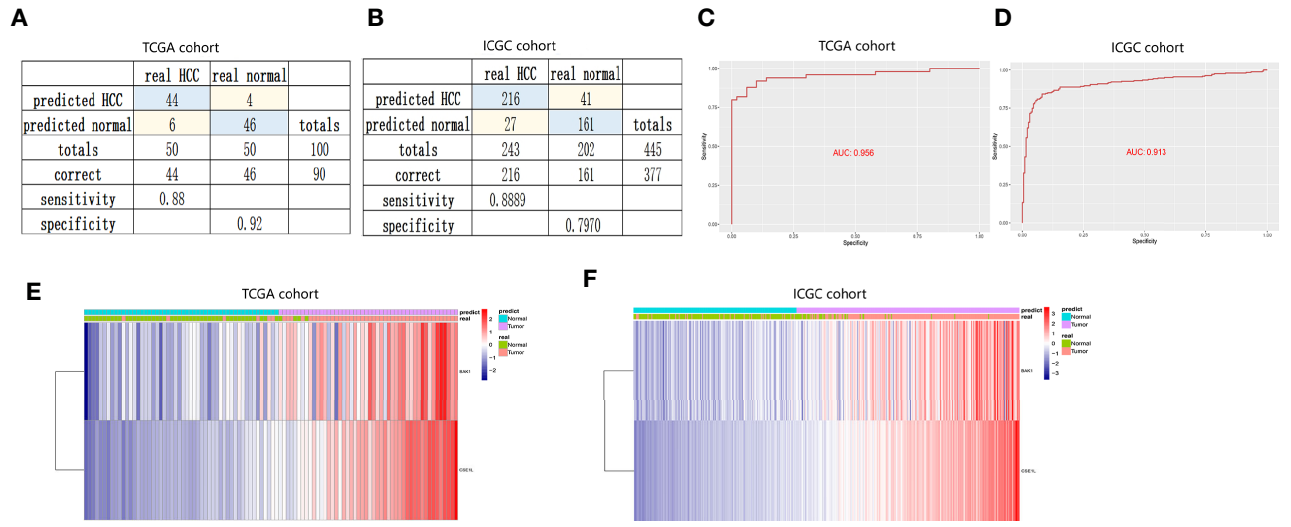


FIGURE 7 | Construction and validation of diagnostic model to distinguish HCC from normal samples. **(A, B)** Confusion matrices of binary data showed high specificity and sensitivity of diagnostic model in the training cohort (TCGA cohort) and validation cohort (ICGC cohort). **(C, D)** Cox analysis and ROC curve predicted the good efficiency of diagnostic model both in TCGA cohort and ICGC cohort. **(E, F)** Unsupervised hierarchical clustering showed the expression of BAK1 and CSE1L in HCC and normal samples.

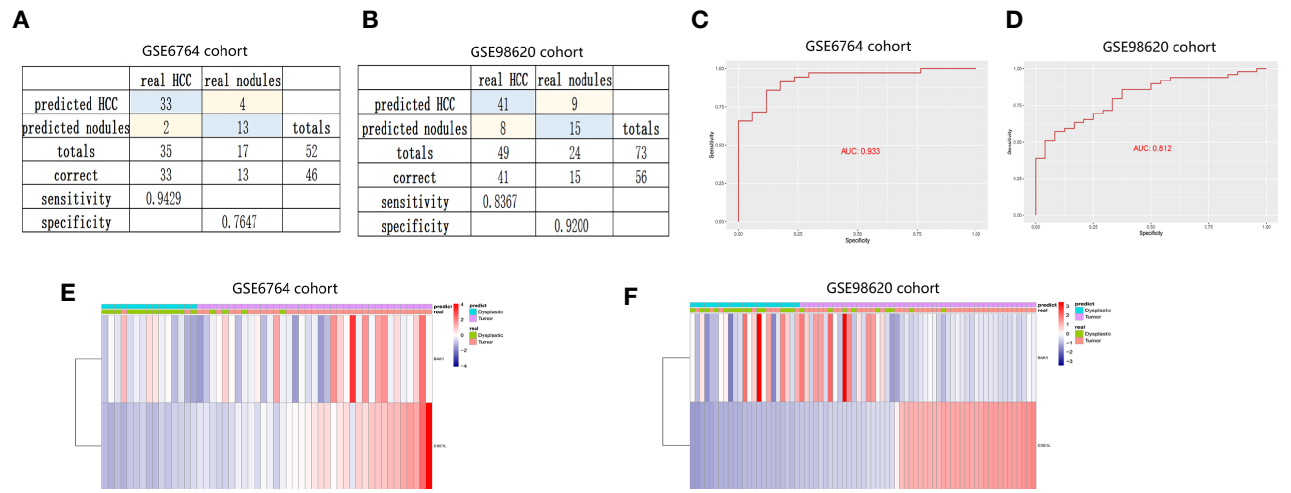


FIGURE 8 | Establishment and validation of diagnostic model to distinguish HCC from dysplastic nodules. **(A, B)** Confusion matrices of binary data indicated high specificity and sensitivity of model to identify HCC and dysplastic nodules in training cohort (GSE6764 cohort) and validation cohort (GSE98620 cohort). **(C, D)** Cox analysis and ROC curve predicted the good efficiency of diagnostic model both in GSE98620 cohort and GSE6764 cohort. **(E, F)** Unsupervised hierarchical clustering indicated the level of BAK1 and CSE1L in HCC and dysplastic nodules.

association between apoptotic genes and the diagnosis and prognosis of HCC through high-throughput sequencing of biomarkers.

In this study, we used high-throughput expression profiling of apoptotic genes to analyze the ability of apoptotic genes to predict prognosis in HCC patients. By comparing the differences in gene expression between HCC and adjacent non-tumor liver tissues in the TCGA database, we screened for differentially expressed apoptotic genes. Subsequently, univariate and multivariate Cox regression analyses were used to identify two apoptotic genes that were significantly associated with HCC prognosis: BAK1 and CSE1L. We thus constructed an apoptosis-related prognostic signature to predict HCC prognosis and survival. The K-M curve and ROC curve indicated that the two apoptotic genes had good prognostic performance. Further validation of the predictive accuracy of the prognostic signature

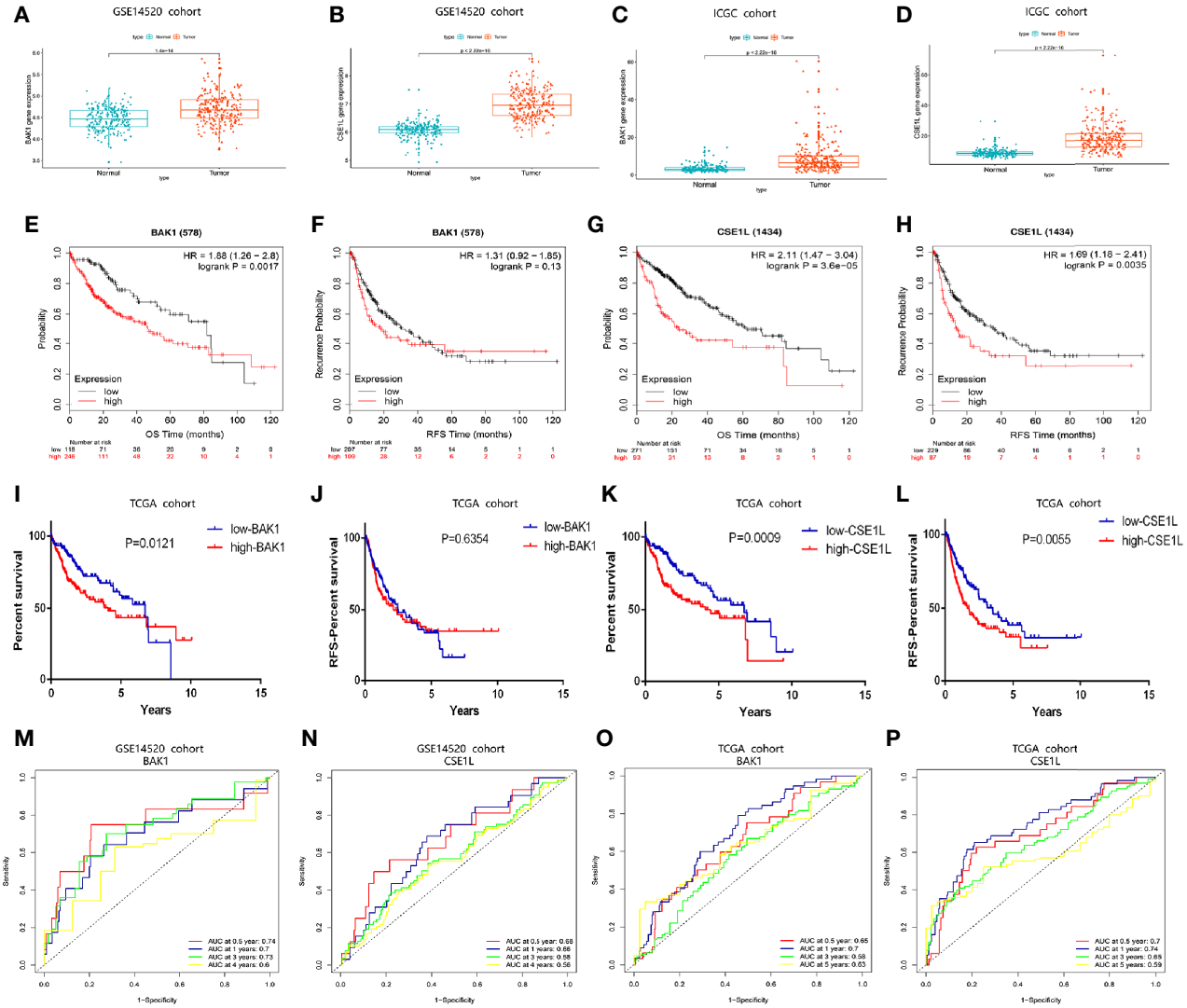


FIGURE 9 | The expression patterns of BAK1 and CSE1L and their predictive power for OS and RFS. **(A, B)** Box plot indicating higher expression of BAK1 and CSE1L in HCC tissues (n = 243) than those in normal liver tissues (n = 241) in GSE14520. **(C, D)** Box plot indicating that expression of BAK1 and CSE1L in HCC tissues (n = 243) were significantly higher than those in normal liver tissues (n = 202) in ICGC. **(E–H)** Survival analysis of BAK1 and CSE1L in Kaplan-Meier Plotte database. **(I, J)** K-M curves indicating the predictive power of different expression of BAK1 for OS and RFS. **(K, L)** K-M curves indicating the predictive power of different expression of CSE1L for OS and RFS. **(M, N)** Time-dependent ROC curve analysis of the two apoptosis genes in the GSE14520 cohort. **(O, P)** Time-dependent ROC curve analysis of the two apoptosis genes in the TCGA cohort.

was performed through GEO, and the expression levels of those two apoptotic genes were analyzed in GSE14520 cohort and ICGC cohort. We also demonstrated that the predictive prognostic ability of this apoptotic-related prognostic signature was independent of other clinical features of HCC by univariate and multivariate Cox regression analyses ($P < 0.05$). A nomogram can predict the probability of a clinical outcome (tumor recurrence or survival) for a given individual to help the clinician make clinical decisions and develop an individualized treatment plan for the patient (27). In this study, we integrated three independent prognostic factors of HCC identified by multivariate Cox regression analysis to construct a nomogram to accurately predict the prognosis and OS of HCC patients. The calibration curve indicated that the actual OS was closely related to the predicted survival rate, suggesting that the nomogram had good prognostic value.

Dysplastic nodules are premalignant neoplastic nodules usually occurred in cirrhotic livers, which divided into nodules (RNs), low grade dysplastic nodules (LGDN), high grade dysplastic nodules (HGDN), and finally small HCC (less than 2 cm). Identification of HCC during early stage is more likely to obtain longer survival time. Currently, the combination of

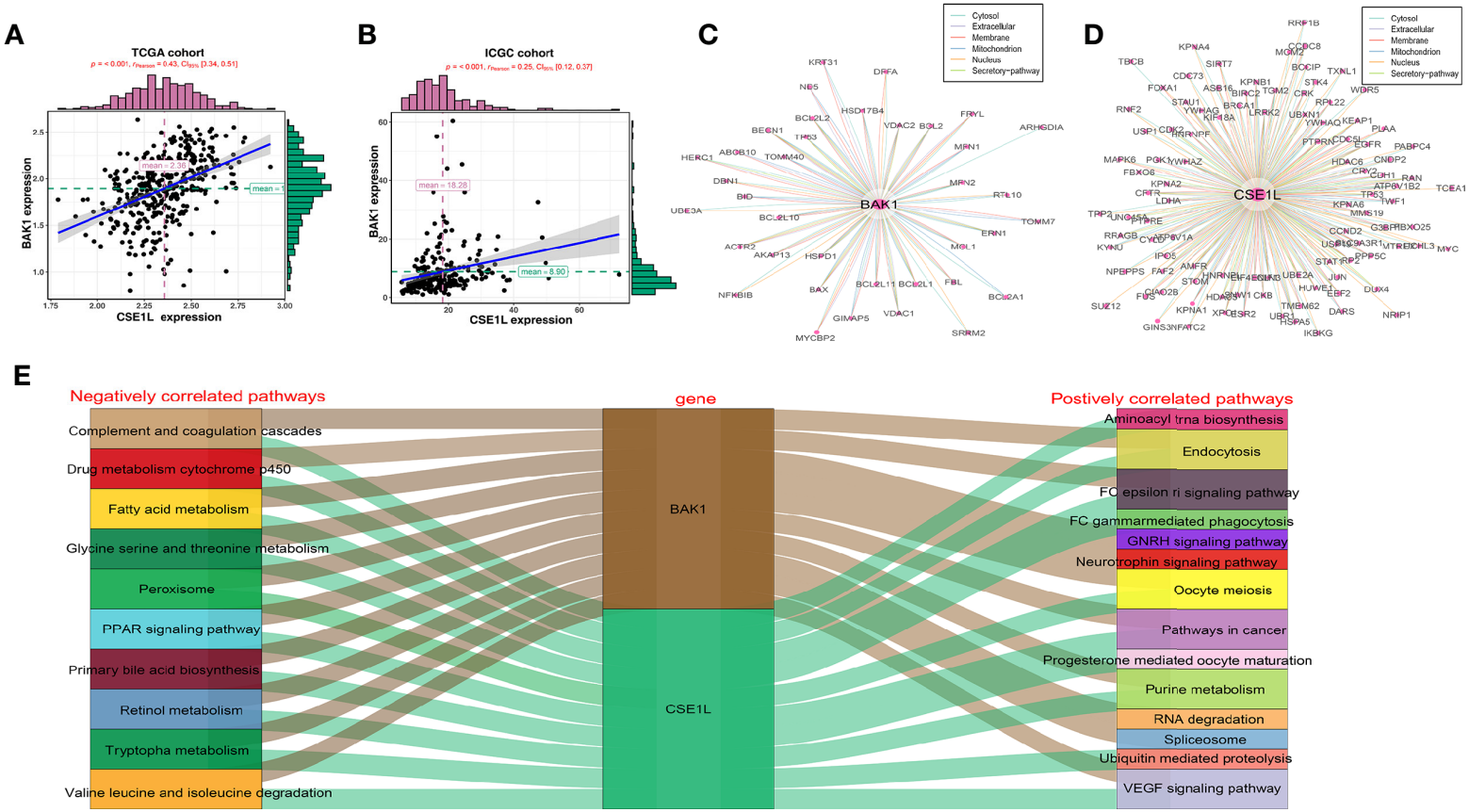


FIGURE 10 | The potential regulatory mechanism of BAK1 and CES1L in the HCC. **(A, B)** Dependence Analysis showed BAK1 expression had positive correlation with CSE1L in the HCC in TCGA cohort and ICGC cohort. **(C, D)** PPI network showed the proteins interacted with BAK1 and CSE1L. **(E)** GSEA analysis indicated BAK1 and CSE1L regulated the underlying signaling pathways in the HCC and the top 10 signaling pathways were showed by Sankey diagram.

radiological examination and pathological diagnosis is mainly used for early diagnosis of HCC, however, small nodules smaller than 2 cm are sometimes difficult to be characterized by pathology or radiology and are likely to be missed (28). The discrimination of dysplastic nodules is not easy based on uncertain clinical feature (29–31), since there are also morphological criteria used to define early HCC in well-differentiated dysplastic nodules (32). Sometimes it is difficult to distinguish the morphological criteria between dysplastic nodules and well-differentiated HCC. Base on the two apoptotic genes, we also constructed two diagnostic models to distinguish HCC from normal samples and dysplastic nodules. In the study, Cox analysis and ROC curve indicated the diagnostic model have good efficiency with high specificity and sensitivity. The great diagnostic performance of the diagnostic model may help detect early small HCC missed by imaging and pathological diagnosis.

CSE1L (chromosome-isolated protein), also known as CAS (apoptosis-susceptible protein), is a multifunctional protein involved in multiple physiological and pathological processes, such as cell survival, apoptosis, nuclear cytoplasmic migration, microencapsulation, and hyperplasia and migration of cancer cells (33). CSE1L/CAS is overexpressed in multiple types of cancers, and many pathological reports have revealed that the expression level of CSE1L is closely associated with cancer proliferation and metastasis (34, 35). BAK1 (BCL2 antagonist/killer 1) is a key regulator of mitochondria-mediated apoptosis (36). The protein encoded by BAK1 belongs to the BCL2 protein family. There has been no research on the association between BAK1 and CSE1L in the diagnosis and prognosis of HCC. BAK1 and CSE1L inhibition suppressed HCC cells proliferation and promoted HCC cells apoptosis. BAK1 and CSE1L may be regard as oncogenic regulators of HCC. Correlation analysis indicated the two genes had positive correlation in HCC. BAK1 and CSE1L also interacted with the same genes including P53 and they regulated same signaling pathway including pathway in cancer, PPAR signaling pathways, etc. That means BAK1 and CSE1L may played synergistic role in the progression and development in the HCC.

Inevitably, our current research has some limitations. First, our study is based on RNA-sequencing data from the public dataset TCGA, and without prospective testing in clinical trials, the predictive performance of the apoptosis-related prognostic signature and nomogram requires further validation in multicenter clinical trials and prospective studies. In addition, the oncogenic effect of BAK1 and CSE1L of HCC need to be elucidated and the underlying mechanism remain to be further explored.

CONCLUSION

In conclusion, we identified the prognostic signature of two apoptosis-related genes as a reliable tool for more accurate prediction of diagnosis and prognosis in HCC patients, and the diagnostic models to distinguish HCC from normal samples/proliferative nodules (≦2cm) with high specificity and sensitivity, which may help doctors to develop more effective early diagnostic and individualized treatment strategies.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study; these can be found in The Cancer Genome Atlas (https://www.tcga.org/) and Deathbase (http://www.deathbase.org/).

AUTHOR CONTRIBUTIONS

JJ and MX conceived and designed the experiments. BT and JZ performed the experiments. XL, MM, QW, and NZ analyzed the data. JL, KF, LZ, and SF contributed analysis tools. BT and JZ wrote the paper. JJ and MX edited the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by National Natural Science Foundation of China (Nos. 81803778), and The Key Research and development Project of Zhejiang Province (No. 2018C0302), and Public Welfare Project of Zhejiang Province (Nos. LGF19H180010, LGD19H160002, and LGF19H180009), and Medical and Health Care Key Project of Zhejiang Province (Nos. 2016146810, 2018KY197, and 2018KY932).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.519180/full#supplementary-material

Supplementary Figure 1 | Functional enrichment analysis of 14 differentially expressed apoptotic genes between HCC tissues and non-cancerous tissues. **(A)** The GO function enrichment analysis of the 14 apoptosis genes. **(B)** The KEGG pathway concentrating on the 14 apoptosis genes.

Supplementary Figure 2 | Distribution of risk score in HCC patients and the survival status of the patients. **(A, C, E)** HCC patients were divided into high-risk group and low-risk group according to the level of risk index in the TCGA cohort **(A)**, GSE14520 cohort **(C)**, and ICGC cohort. **(B, D, F)** The heat map shows the relationship between different risk scores and the expression levels of BAK1 and CSE1L in the TCGA cohort **(B)**, GSE14520 cohort **(D)**, and ICGC cohort **(F)**.

Supplementary Table 1 | Apoptosis genes extracted from the Deathbase.

Supplementary Table 2 | Univariate and multivariate Cox regression analyses of the prognostic signature and clinical features related to OS in HCC patients.

Supplementary Table 3 | Underlying signal pathway in HCC regulated by BAK1 determined by GSEA analysis

Supplementary Table 4 | Underlying signal pathway in HCC regulated by CSE1L determined by GSEA analysis

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* (2007) 132:2557–76. doi: 10.1053/j.gastro.2007.04.061
3. Mohammadian M, Soroush A, Mohammadian-Hafshejani A, Towhidi F, Hadadian F, Salehiniya H. Incidence and Mortality of Liver Cancer and Their Relationship with Development in Asia. *Asian Pac J Cancer Prev* (2016) 17:2041–7. doi: 10.7314/APJCP.2016.17.4.2041
4. Grandhi MS, Kim AK, Ronnekleiv-Kelly SM, Kamel IR, Ghasebeh MA, Pawlik TM. Hepatocellular carcinoma: From diagnosis to treatment. *Surg Oncol* (2016) 25:74–85. doi: 10.1016/j.suronc.2016.03.002
5. Schlachterman A, Craft WW Jr., Hilgenfeldt E, Mitra A, Cabrera R. Current and future treatments for hepatocellular carcinoma. *World J Gastroenterol* (2015) 21:8478–91. doi: 10.3748/wjg.v21.i28.8478
6. Katanoda K, Matsuda T. Five-year relative survival rate of liver cancer in the USA, Europe and Japan. *Jpn J Clin Oncol* (2014) 44:302–3. doi: 10.1093/jjco/hyu025
7. Naito S, Imamura H, Tukada A, Matsuyama Y, Yoshimoto J, Sugo H, et al. Postoperative recurrence pattern and prognosis of patients with hepatocellular carcinoma, with particular reference to the hepatitis viral infection status. *Liver Int* (2014) 34:802–13. doi: 10.1111/liv.12447
8. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nat Rev Cancer* (2006) 6:674–87. doi: 10.1038/nrc1934
9. Elmore S. Apoptosis: a review of programmed cell death. *Toxicol Pathol* (2007) 35:495–516. doi: 10.1080/01926230701320337
10. Tower J. Programmed cell death in aging. *Ageing Res Rev* (2015) 23:90–100. doi: 10.1016/j.arr.2015.04.002
11. Renehan AG, Booth C, Potten CS. What is apoptosis, and why is it important? *BMJ (Clinical Res ed)* (2001) 322:1536–8. doi: 10.1136/bmj.322.7301.1536
12. Fadeel B, Orrenius S. Apoptosis: a basic biological phenomenon with wide-ranging implications in human disease. *J Internal Med* (2005) 258:479–517. doi: 10.1111/j.1365-2796.2005.01570.x
13. Su C. Survivin in survival of hepatocellular carcinoma. *Cancer Lett* (2016) 379:184–90. doi: 10.1016/j.canlet.2015.06.016
14. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* (2006) 26:565–74. doi: 10.1177/0272989X06295361
15. Zhu J, Tang B, Li J, Shi Y, Chen M, Lv X, et al. Identification and validation of the angiogenic genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma. *Aging (Albany NY)* (2020) 12(9):7848–73. doi: 10.18632/aging.103107
16. Menyhart O, Nagy Á, Győrffy B. Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. *R Soc Open Sci* (2018) 5:181006. doi: 10.1098/rsos.181006
17. Klungboonkrong V, Das D, McLennan G. Molecular Mechanisms and Targets of Therapy for Hepatocellular Carcinoma. *J Vasc Interv Radiol* (2017) 28:949–55. doi: 10.1016/j.jvir.2017.03.002
18. Cong WM, Bu H, Chen J, Dong H, Zhu YY, Feng LH, et al. Practice guidelines for the pathological diagnosis of primary liver cancer: 2015 update. *World J Gastroenterol* (2016) 22:9279–87. doi: 10.3748/wjg.v22.i42.9279
19. Chou R, Cuevas C, Fu R, Devine B, Wasson N, Ginsburg A, et al. Imaging Techniques for the Diagnosis of Hepatocellular Carcinoma: A Systematic Review and Meta-analysis. *Ann Internal Med* (2015) 162:697–711. doi: 10.7326/M14-2509
20. Jiang HY, Chen J, Xia CC, Cao LK, Duan T, Song B. Noninvasive imaging of hepatocellular carcinoma: From diagnosis to prognosis. *World J Gastroenterol* (2018) 24:2348–62. doi: 10.3748/wjg.v24.i22.2348
21. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet (London England)* (2012) 379:1245–55. doi: 10.1016/S0140-6736(11)61347-0
22. Chen SS, Yu KK, Ling QX, Huang C, Li N, Zheng JM, et al. The combination of three molecular markers can be a valuable predictive tool for the prognosis of hepatocellular carcinoma patients. *Sci Rep* (2016) 6:24582. doi: 10.1038/srep24582
23. Li L, Lei Q, Zhang S, Kong L, Qin B. Screening and identification of key biomarkers in hepatocellular carcinoma: Evidence from bioinformatic analysis. *Oncol Rep* (2017) 38:2607–18. doi: 10.3892/or.2017.5946
24. Cong WM, Wu MC. New insights into molecular diagnostic pathology of primary liver cancer: Advances and challenges. *Cancer Lett* (2015) 368:14–9. doi: 10.1016/j.canlet.2015.07.043
25. Pistritto G, Trisciuoglio D, Ceci C, Garufi A, D'Orazi G. Apoptosis as anticancer mechanism: function and dysfunction of its modulators and targeted therapeutic strategies. *Aging* (2016) 8:603–19. doi: 10.18632/aging.100934
26. Wong RS. Apoptosis in cancer: from pathogenesis to treatment. *J Exp Clin Cancer Res* (2011) 30:87. doi: 10.1186/1756-9966-30-87
27. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* (2015) 16:e173–80. doi: 10.1016/S1470-2045(14)71116-7
28. Kojiro M. Focus on dysplastic nodules and early hepatocellular carcinoma: an Eastern point of view. *Liver Transpl* (2004) 10:S3–8. doi: 10.1002/lt.20042
29. The International Consensus Group for Hepatocellular Neoplasia. Pathologic diagnosis of early hepatocellular carcinoma: a report of the international consensus group for hepatocellular neoplasia. *Hepatology (Baltimore Md)* (2009) 49:658–64. doi: 10.1002/hep.22709
30. Hytiroglou P, Park YN, Krinsky G, Theise ND. Hepatic precancerous lesions and small hepatocellular carcinoma. *Gastroenterol Clin North Am* (2007) 36:867–87. doi: 10.1016/j.gtc.2007.08.010
31. Peng C, Kaščáková S, Chiappini F, Olaya N, Sandt C, Yousef I, et al. Discrimination of cirrhotic nodules, dysplastic lesions and hepatocellular carcinoma by their vibrational signature. *J Trans Med* (2016) 14:9. doi: 10.1186/s12967-016-0763-6
32. International Working Party Terminology of nodular hepatocellular lesions. *Hepatology (Baltimore Md)* (1995) 22:983–93. doi: 10.1016/0270-9139(95)90324-0
33. Jiang MC. CAS (CSE1L) signaling pathway in tumor progression and its potential as a biomarker and target for targeted therapy. *Tumour Biol* (2016) 37:13077–90. doi: 10.1007/s13277-016-5301-x
34. Tai CJ, Hsu CH, Shen SC, Lee WR, Jiang MC. Cellular apoptosis susceptibility (CSE1L/CAS) protein in cancer metastasis and chemotherapeutic drug-induced apoptosis. *J Exp Clin Cancer Res* (2010) 29:110. doi: 10.1186/1756-9966-29-110
35. Tung MC, Tsai CS, Tung JN, Tsao TY, Chen HC, Yeh KT, et al. Higher prevalence of secretory CSE1L/CAS in sera of patients with metastatic cancer. *Cancer Epidemiol Biomarkers Prev* (2009) 18:1570–7. doi: 10.1158/1055-9965.EPI-08-0948
36. Shimizu S, Narita M, Tsujimoto Y. Bcl-2 family proteins regulate the release of apoptogenic cytochrome c by the mitochondrial channel VDAC. *Nature* (1999) 399:483–7. doi: 10.1038/20959

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 Zhu, Tang, Lv, Meng, Weng, Zhang, Li, Fan, Zheng, Fang, Xu and Ji. 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 Oncology

REVIEW
published: 23 February 2021
doi: 10.3389/fonc.2021.559822

Application of Advanced Mass Spectrometry-Based Proteomics to Study Hypoxia Driven Cancer Progression

Arada Vinaiphat[1], *Jee Keem Low*[2], *Kheng Wei Yeoh*[3], *Wee Joo Chng*[4] *and Siu Kwan Sze*[1]*

[1] School of Biological Sciences, Nanyang Technological University, Singapore, Singapore, [2] Department of Surgery, Tan Tock Seng Hospital, Singapore, Singapore, [3] Department of Radiation Oncology, National Cancer Centre Singapore, Singapore, Singapore, [4] Department of Hematology-Oncology, National University Cancer Institute, National University Health System, Singapore, Singapore

OPEN ACCESS

Edited by:
Bhaswati Chatterjee,
National Institute of Pharmaceutical Education and Research, India

Reviewed by:
Paul Townsend,
The University of Manchester, United Kingdom
Ornella Cominetti,
Nestle Institute of Health Sciences (NIHS), Switzerland

***Correspondence:**
Siu Kwan Sze
sksze@ntu.edu.sg

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 07 May 2020
Accepted: 07 January 2021
Published: 23 February 2021

Citation:
Vinaiphat A, Low JK, Yeoh KW, Chng WJ and Sze SK (2021) Application of Advanced Mass Spectrometry-Based Proteomics to Study Hypoxia Driven Cancer Progression. Front. Oncol. 11:559822. doi: 10.3389/fonc.2021.559822

Cancer is one of the largest contributors to the burden of chronic disease in the world and is the second leading cause of death globally. It is associated with episodes of low-oxygen stress (hypoxia or ischemia/reperfusion) that promotes cancer progression and therapeutic resistance. Efforts have been made in the past using traditional proteomic approaches to decipher oxygen deprivation stress-related mechanisms of the disease initiation and progression and to identify key proteins as a therapeutic target for the treatment and prevention. Despite the potential benefits of proteomic in translational research for the discovery of new drugs, the therapeutic outcome with this approach has not met expectations in clinical trials. This is mainly due to the disease complexity which possess a multifaceted molecular pathology. Therefore, novel strategies to identify and characterize clinically important sets of modulators and molecular events for multi-target drug discovery are needed. Here, we review important past and current studies on proteomics in cancer with an emphasis on recent pioneered labeling approaches in mass spectrometry (MS)-based systematic quantitative analysis to improve clinical success. We also discuss the results of the selected innovative publications that integrate advanced proteomic technologies (*e.g.* MALDI-MSI, pSILAC/SILAC/iTRAQ/TMT-LC-MS/MS, MRM-MS) for comprehensive analysis of proteome dynamics in different biosystems, including cell type, cell species, and subcellular proteome (*i.e.* secretome and chromatome). Finally, we discuss the future direction and challenges in the application of these technological advancements in mass spectrometry within the context of cancer and hypoxia.

Keywords: cancer, hypoxia, proteomics, mass spectrometry, SILAC, iTRAQ, TMT, MRM

INTRODUCTION

The global burden of cancer is substantial with millions of lives being lost each year (1, 2). In 2018, there were 18.1 million new diagnoses and 9.6 million deaths due to it (2). In addition, it also inflicts considerable morbidity to those affected both physically and psychologically. Advances in medical sciences have been tremendous with improvement in disease prevention, diagnosis, screening and treatment resulting in a decline in death rates for certain cancers (1, 2). For example, the development of the human papillomavirus (HPV) vaccine and the Papanicolaou (PAP) smear test for early detection of cervical cancer has decreased the death rates in the last half-century, and may even eliminate this cancer in the coming years (3–5).

Despite the improved understanding of cancer pathology and the remarkable leaps made in cancer treatment with a move towards personalized approaches, there remains challenges in current treatment approaches to eradicate and control cancer (6). Cancer is a complex disease that can no longer be understood as a simple enumeration of identical cancerous cells. Genetic diversity between different types of tumors and different patients makes it exceedingly difficult to eliminate all cancer cell types using a single therapeutic strategy (7–9). More importantly, recent studies using single-cancer-cell DNA sequencing to elucidate genetic diversification have revealed the heterogeneity among cancer cell populations (10–13). Heterogeneous cancer cells within a tumor, when subjected to selective pressures (Darwinian's law of natural selection) either from treatment (chemotherapy, radiotherapy and biologicals) or unfavorable microenvironmental conditions such as hypoxia, results in the adaptation and selection of subclones with improved survival and malignant potential; with ensuing therapeutic resistance clinically (7, 9).

To overcome the limitations of current therapeutic approaches, we need to identify molecular signatures of malignant cell transformation and understand the carcinogenic processes involved. This will enable us to explicitly pinpoint precise targets for intervention as well as identifying novel biomarkers to monitor cancer progression and treatment response. While most advances in treatment have been based on genetic determinants (6), how these genetic changes translate to proteomic and phenotypic characteristics has not been well characterized; thus opening up opportunities for research in the field of proteomics (14). We will examine some of these advances with a focus on hypoxia which plays a central role in cancer progression and treatment resistance.

HYPOXIA-DRIVEN CANCER AND HYPOXIA-TARGETED THERAPY

As aforementioned, the genetic diversity across the same and different types of tumors imposes challenges in finding the therapeutic targets that are unique to all cancerous cells, thus limiting the effectiveness of most cancer drugs. To worsen the situation, microenvironment of the tumor itself, in fact, is another formidable feature that escalates the complexity of cancer dynamic (8, 15, 16).

In addition to the complex dynamic interaction between tumor and stromal cells in the tumor microenvironment, changes in the availability of oxygen has been found to cause alteration in epigenomic, genomic, and proteomic profiles in cancer cells. In normal physiological condition, the oxygen tension of the cell environment is between 5 and 10%, whereas hypoxia is a condition in which there is a deprivation in oxygen level around 1–2% or below (17). The role of hypoxia in carcinogenesis and therapy resistance has been recognized for at least half a century (18). At the initial phase of solid tumor growth, the microenvironment of tumors with a diameter beyond 1 mm become hypoxic due to inadequate oxygen supply (**Figure 1**) (19).

Under hypoxic conditions, tumor cells undergo specific gene translational changes, following the intracellular signaling pathways alteration. One of the well-studied modulators of the cellular response to hypoxia is the hypoxia-inducible factor (HIF) transcription factor family (20–23) (**Figure 2**). HIF contains a heterodimer of an oxygen labile subunit HIF-α and an oxygen-insensitive HIF-1β (24). In normoxia, HIF-α are hydroxylated by Prolyl-Hydroxylase (PHD) and Factor Inhibiting HIF (FIH-1), oxygen sensor enzymes, leading to rapid proteasomal degradation of HIF-α subunits. Contrarily, the deprivation of oxygen will lead to the accumulation of HIF-α and the translocation into the nucleus, resulting in the activation of genes involved in adaptive cellular changes as a response to low oxygen stress (25). It is noted that HIF mediated gene expression can also be achieved by hypoxia-independent activation of reactive oxygen species, nitric oxide, cytokines, G protein-coupled receptors, toll-like receptors and alarmins receptors, involving pathways such as PI3K/AKT/mTOR (26, 27), NF-$k\beta$ (28), p38 and ERK (29). As a result, these pathways are shown to regulate cells apoptosis (*via* BNIP-3 and p53), metabolism (*via* GLUT-1 and GSK), metastasis (*via* EMT, CXCR4, E-cad, CAIX, LOX, and MMPs) and vessel formation (via VEGF, SDF-1, Ang-2, MMPs) (17). Such changes as a response to hypoxia have been demonstrated to ultimately drive tumor progression, enhance aggression, and promote metastatic phenotypes (15, 17).

In addition, it is worth to clarify that an induction of HIF pathway with resultant angiogenesis does not lead to normoxia recovery (30). Although the newly formed vessels can provide rapid-growing cancer cells with nutrients and oxygen for survival, the hypoxia-induced neovasculogenesis often results in an abnormal and leaky vascular network, causing irregular and sluggish blood flow (**Figure 1**). Consequently, this dysfunctional vascularization will eventually fail to provide enough blood supply for the fast-expanding cancer population (31), and vicious cycle of hypoxia condition begins (32).

As previously discussed, hypoxia can establish new cell phenotypes through induction of various signaling pathways, affecting clinical responses to therapy. Tumor hypoxia has been demonstrated to increase the expression of anti-apoptotic Bcl-2

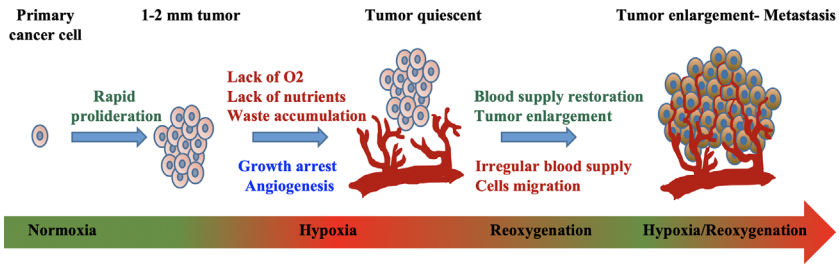


FIGURE 1 | Mechanism of hypoxia-induced tumor development. Exponential cellular proliferation results in the formation of tumor. Impenetrable blood vessel that supplies oxygen and nutrients causes hypoxic environment where waste accumulates and growth arrest. Tumor cell adaptation such as the alteration in cellular metabolism, angiogenesis, vasculogenesis, and metastasis lead to irregular blood supply, all of which further promote tumor progression. Modified from "Quantitative profiling of chromatome dynamics reveals a novel role for HP1BP3 in hypoxia-induced oncogenesis," by B. Dutta et al. (19), Molecular & Cellular Proteomics, 13(12), p. 3236–3249. Copyright 2014 by the American Society for Biochemistry and Molecular Biology, Inc. Reprinted with permission.

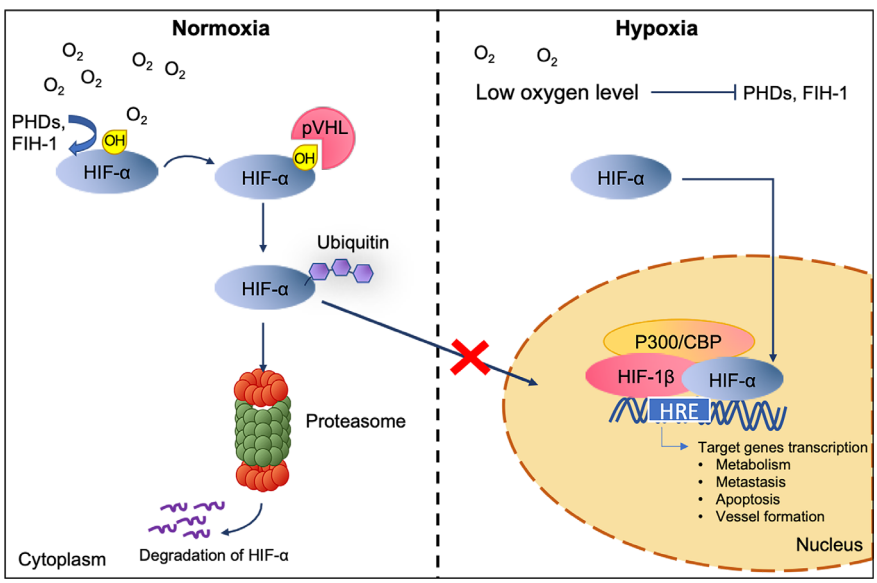


FIGURE 2 | HIF interaction and degradation pathway in normoxic and hypoxic conditions. The level of HIF-α is maintained by the recognition of the von-Hippel Lindau tumor suppressor (pVHL) to the hydroxylated HIF-α subunit for ubiquitylation and subsequently proteasomal degradation under normoxic. Under low oxygen level, untargeted HIF-α subunits accumulate and translocate to the nucleus. Once inside, dimerization of HIF-α with HIF-1β and its co-activators p300/CBP activate the target genes transcription (20).

and Bcl-xL and decrease pro-apoptotic Bcl-2 such as Bax, Bad and Bid. Failure to activate stress-mediated apoptosis in tumor hypoxia 'primed' tumor cells to become resistant to cancer treatment (33, 34). Similarly, hypoxia-induced autophagy activation or cell cycle arrest at G1 phase in cancer cells as an adaptation to stress lead to cell survival under cancer therapy (34). The resistance of cancer cells to treatment has challenged the medical industry as many conventional drugs and radiotherapy aim to induce cancer cells apoptosis *via* endogenous mechanisms (35, 36). Example of current drugs use in cancer that have shown to be less effective in hypoxic regions in tumors are 5-Fluorouracil (37), Actinomycin D (37–39), Sorafenib (40) and Bleomycin (37–39).

Undoubtedly, hypoxic regions in tumors have become one of the mechanistic conduits that promote cancer progression. However, tumor hypoxia is exploitable in cancer treatment. For instance, disrupting molecular mechanisms underlying hypoxia-induced adaptation may be the key to the future development of effective therapies (41, 42). In fact, hypoxia-selective cytotoxins and drugs are emerging as a new strategy to treat different types of cancer (43). Examples of drugs that directly inhibits HIF activity are the well-established rapamycin (inhibits mRNA/protein expression) (44), Aminoflavone (inhibits mRNA/protein expression (45), Acriflavine (inhibits HIF-α/HIF-1β dimerization) (46) and Bortezomib (inhibits transcriptional activity) (47). Furthermore, hypoxia prodrugs

that are inactive in normoxic but are toxic under hypoxic condition have been developed to selectively target hypoxic tumor cells (48). However, the results of many of these drugs in preclinical and clinical trials do not show maximum therapeutic effect. This may be due to the lack of 'personalized' and tumor-specific target (43). In this respect, discovering potential hypoxic-associated tumor development pathways may lead to a successful hypoxia-targeted therapy.

MASS SPECTROMETRY-BASED PROTEOMICS APPROACHES FOR HYPOXIA-ASSOCIATED CANCER RESEARCH

Mass spectrometry (MS) for proteomics is a powerful analytical technique that has the potential to revolutionize biomedical research. The development of different MS ionization techniques [electronspray ionization; ESI (49) and matrix-assisted laser desorption/ionization; MALDI (50)], mass analyzers (quadrupole; quadruple ion trap, QIT/linear ion trap, LIT or LTQ; time-of-flight, TOF; Fourier-transform ion cyclotron resonance, FTICR, and Orbitrap) (51) and fragmentation methods in tandem mass spectrometry (MS/MS) [collision-activated dissociation, CAD (52)/collision-induced dissociation, CID (53); electron ionization dissociation, EID (54); electron capture dissociation (55, 56)/ electron-transfer dissociation, ETD (57)] has presented an increasing power of both protein identification and analysis at different dimensions (*e.g.* protein expression, structure, interaction, modification, *etc.*); and therefore, allows greater understanding of complex biological processes and diseases (58–60).

Like many other research fields, cancer hypoxia has greatly benefited from its recent development in MS-based proteomics for the identification of key players and the underlying pathological mechanisms through protein profiling in complex biological samples. In such studies, differential proteomes across various types of cells/tissue/clinical samples in association of hypoxic conditions are analyzed. Proteins that are repeatedly identified across different samples suggest their likelihood of having an active role in hypoxia regulation and thus may lead to the generation of novel targeted cancer therapies (**Table 1**) (76, 77). On the other hand, a protein (or a set of proteins) of high confidence that is (are) exclusively identified in specific samples could act as candidate biomarkers for clinical use (78–81). In the following sections, we collectively describe the successful application of advance quantitative proteomics approaches coupled with different MS analytical platforms and experimental strategies which have been undertaken to overcome present limitations, focusing on studies published within the last ten years (82).This will allow us to gain valuable insights for meaningful translational research in the field of hypoxia-associated cancer pathology.

ADVANCE MS-BASED STRATEGIES

Although a multitude of hypoxia mediators and biomarkers have been identified and quantified using label-free relative quantification, there has been a remarkable transition from

TABLE 1 | Comprehensive list of the potential therapeutic targets from labeled and/or Targeted quantitative proteomic approach in hypoxic cancer cells.

Method	Potential biomarkers for cancer diagnosis	Cell line	Sample type	Reference
Metabolic labeling				
SILAC	GM3 synthase	HeLa	WCL	(61)
	Subunit B of respiratory complex II			
	Formin-2	NARF2	WCL	(62)
	Retinoic acid-induced protein 3	SW620	WCL	(63)
	Histone demethylase jumonji domain containing protein 2C	HeLa	WCL-pulldown	(64)
	Lysyl oxidase	MDA-MB-231	Secretome	(65)
pSILAC	PHD finger protein 14	A431	WCL	(66)
	Endoplasmic reticulum disulphide oxidase 1α	MIAPaCa-2	WCL	(67)
	Hypoxia-induced angiogenesis regulator	NCBP2-AS2	Secretome	(68)
Chemical labeling				
iTRAQ	Ku70/Ku80 dimer	A431	WCL	(69)
	Integrin $\alpha 3$	A431	WCL	(70)
	Calmodulin-1	A431	WCL	(71)
	Calumenin			
	Reticulocalbin-1			
	Heterochromatin protein 1-binding protein 3	A431	Chromatome	(19, 72)
TMT	Vitamin B_{12} transporter protein TCN2	U87	WCL	(73)
	Colony stimulating factor-1	B16-F0	Exosome	(74)
	Ferritin heavy/light chain			
	Lysyl oxidase	U87	Exosome	(75)
	A disintegrin and metalloproteinase with thrombospondin motifs 1			
	Vascular endothelial growth factor A			
Targeted				
MRM	Ku70/Ku80 dimer	A431	WCL	(69)

WCL, Whole cell lysate.

static proteome analysis to elucidate diverse aspects of the dynamic proteome through new techniques such as mass spectrometry imaging (MSI) for spatial distribution of molecules (83) or multiple reaction monitoring (MRM) and stable heavy isotope labeling combined with MS for targeted (84) and absolute quantitation (85). The different uses of advance multiplexed quantitative techniques, whether carried out independently or in combination, have enabled investigators to quantify precise changes in protein/peptide and PTM abundances in multiple biological matrices simultaneously within a single MS analysis (86–88). A summary of quantitative proteomic analysis is depicted in **Figure 3**. The improvement in robustness, resolution, and specificity has proven useful especially in the identification of true protein interactions (89, 90), detecting the presence of specific subset of proteins/peptides in highly complex samples (biomarkers) (91, 92) and defining accurate relationship between mechanism of signaling pathway and its downstream biological responses (93, 94).

Metabolic Labeling Coupled With MS

Metabolic labeling with ^{15}N is the very first isotope-based labeling technique to be applied for quantitative proteomic analysis (95). After the incorporation of isotopes to live cells through growth medium, the acquired samples can be combined prior to cell lysis, thus minimizing errors that may arise during subsequent multi-step proteomics sample preparation from purification to fractionation steps before MS analysis. The introduction of mass shift to sample peptides with the labeled heavy isotype of nitrogen allows relative quantification unaffected by sample purity or instrument response when comparing spectral intensities of heavy peptide to that of naturally occurring light peptide (^{14}N) in control sample (96, 97). However, this method limits the analysis to two samples per experiment and it is not suitable for organisms with long generation times, and high ^{15}N enriched isotope (>99%) is necessary for less dispersive peptide isotope distribution and accurate quantitation. The later developed method of stable isotope labeling by amino acids in cell culture (SILAC) for whole-proteome MS analysis of multiple samples (up to five) in different conditions has become a popular alternative approach. In this technique, only selected amino acids are being labeled. Typically, arginine and lysine are selected since the subsequence proteolytic cleavage with trypsin during proteomics workflow will result in a pool of peptides that contain at least one labeled amino acid for downstream quantification by MS (96, 98). It is noted that although metabolic labeling requires living cells, recent studies have expanded the use of SILAC to tissue analysis using super-SILAC approach (99–101).

Application of SILAC Labeling and LC–MS/MS for Quantitative Proteome Analysis of Hypoxia-Associated Proteins *In Vitro*

The proteomic analysis based on SILAC and high resolution tandem mass spectrometry such as Orbitrap mass spectrometer is often the method of choice to identify and quantify cancer hypoxia-regulated proteins from complex peptide mixtures. One important work by Yates et al. has demonstrated that the combination of the high-throughput, the high-resolution and high-accuracy capabilities of the Orbitrap mass analyzer enabled four- to five-fold improvements in the number and quality of the peptide ratio measurements compared to low resolution LTQ (102). Using this method, many studies have successfully characterized novel hypoxia-perturbed pathways and mediators in various cell types of human system, including breast cancer (103), cervical carcinoma (HeLa cells) (61, 104), osteosarcoma (62), colorectal tumor (63).

SILAC-based LC–MS/MS approach enabled researchers to descriptively explain the global changes and adaptation in tumor cells under oxygen deprivation (103, 104). In 2014, a comprehensive proteome analysis was conducted on 4T1 cells of breast cancer model to monitor the ratio of differentially regulated proteins under hypoxic conditions (103). SILAC in

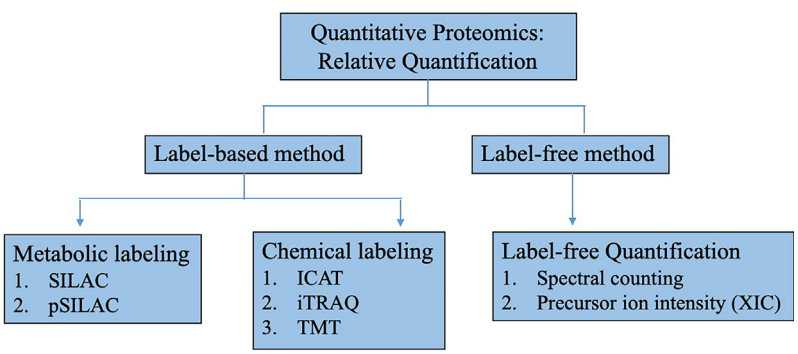


FIGURE 3 | Methods for relative quantitative proteomics research. Relative quantitative proteomics can be classified into label-based or label free methodologies. In label-based method, labeling can be done by metabolically incorporating stable heavy isotope into proteins such as in stable isotope labeling with amino acids in cell culture (SILAC). Protein/peptides can also be chemically labeled with isotopic tag after sample preparation such as isotope-code affinity tag (ICAT), isobaric tags for relative and absolute quantitation (iTRAQ) and tandem mass tags (TMT). For label-free method, quantification is done by tandem mass spectral counting or MS precursor ion intensity through extracted ion chromatogram (XIC).

conjunction with LTQ-Orbitrap has identified a total of 131 proteins with significant quantitative changes, 60 of which were upregulated (*e.g.* glyceraldehyde-3-phosphate dehydrogenase, L-lactate dehydrogenase A chain, and phosphoglycerate kinase) and 71 downregulated (*e.g.* ribosomal proteins and ubiquitin-associated proteins). The classification of altered proteins according to biological process and molecular functions using PANTHER bioinformatic tool (www.pantherdb.org) revealed that hypoxia-associated proteins are mainly involved in metabolic processes, cellular process, binding, and catalytic activities.

In line with proteomic study on breast cancer model, another quantitative proteome analysis conducted by Bousquet et al. in 2015 using HeLa cells as a representative of human cancer system has revealed some similar findings (104). In this study, 125 unique hypoxia-regulated proteins were identified. 72 proteins that were found upregulated are mostly proteins of glycolysis (noted that lactate dehydrogenase A, phosphoglycerate kinase 1 and 2 were also identified; Supplementary Table 1). However, interestingly, most of the 53 downregulated proteins were identified as mitochondrial proteins (*e.g.* mitochondrial ribosomal proteins and mitochondrial translocases—these proteins have not previously been reported as hypoxia-regulated proteins in tumor cells) and a few were citric acid cycle (CAC) related proteins. Again, Gene Ontology based analysis by DAVID (http://david.abcc.ncifcrf.gov/) and PANTHER also revealed that the main biological process affected by hypoxia was metabolism, and that proteins with catalytic and binding activities were shown to dominate the molecular function. Collectively, the upregulation of various glycolytic enzymes, particularly, lactate dehydrogenase, which function to convert pyruvate to lactate, and the downregulation of mitochondrial proteins suggested metabolic shifts of cancer cells from cellular respiration to inefficient glycolysis for ATP needs *via* enhancement of glycolysis and suppression of cellular respiration in the absence of oxygen. Together, the change in metabolic pathways-related proteins under hypoxia condition supports the key difference between normal tissue and cancer metabolism which was first noted since 1920s. These cancer cells undergo a phenomenon termed the Warburg Effect, shifting from oxidative phosphorylation to aerobic glycolysis despite the availability of oxygen (105). In addition, Bousquet et al. (2018) expanded the use of his latter quantitative proteomic data to explain the potential mechanism of N-glycolyl (NeuGc) GM3, a tumor-associated antigen, generation in human cancer cell line induced by hypoxia (61). Specifically, GM3 synthase and subunit B of respiratory complex II (SDHB) that were found upregulated under hypoxia condition may be responsible for the increase in NeuGc incorporation into human cancer cells. Despite the two proteins being potential targets for cancer therapy, further validation of the hypothesis is needed.

Several novel mediators of the hypoxia-induced adaptive response that function to protect cancer cell from apoptosis were successfully identified by SILAC and Orbitrap. In 2013, quantitative analysis of nucleolar proteome alteration following induction of p14ARF, a tumor suppressor gene, was performed on NARF2 osteosarcoma cells (62). The MS analysis identified relative expression ratio of more than thousands of nucleolar proteins, and that Formin-2 (FMN2) was highly induced by ARF. Subsequence analysis using qPCR confirmed that FMN2 expression can also be induced by hypoxia (%1 O_2), and the depletion of FMN2 resulted in apoptosis induction. Similarly, the most recent study conducted by Greenhough et al. in 2018 reported the upregulation of GPRC5A in colorectal tumor cells grown under hypoxia stress (63) despite unchanged levels found in previous omics studies on HeLa cells (SILAC and Orbitrap) (104) and breast cancer cells (SILAC and Orbitrap based LC-MS/MS and MALDI-MSI) (103). This suggested that GPRC5A could be a colorectal cancer-specific hypoxia mediator. Functional analysis of GPRC5A confirmed that GPRC5A promoted hypoxic cancer cells survival *via* the Hippo pathway effector YAP. Collectively, these data highlight FMN2 and GPRC5A as targets for cellular vulnerabilities of cancer cells.

An important co-activator of hypoxia-inducible factor 1α (HIF-1α) was verified using pull-down of SILAC labeled targets coupled with the high-performance quadruple time-of-flight (Q-TOF) mass spectrometer (64). As previously mentioned, hypoxia-inducible factors (HIFs) are activated in solid tumors as a result of hypoxia exposure. The activation of HIF in turn targets genes encoding proteins that are critical in cancer progression (17). In this study, endogenous HIF-1α and HIF-2α-interacting proteins in HeLa cells treated with prolyl hydroxylase inhibitor, dimethyloxalylglycine (DMOG), to induce HIF transcriptional activity were screened using GST fusion protein containing protein binding domain of HIF-1α (residues 531–826) or HIF-2α (450–870) as bait for pulling down interacting partners using quantitative isotope labeling by SILAC. MS identified a total of 146 proteins as HIF-1α and HIF-2α-interacting proteins, with 44 proteins as specific HIF-1α-interacting partners and 42 proteins as specific HIF-2α-interacting partners. The ratio of peak intensities obtained from HIF-1α (heavy peaks) or HIF-2α (medium peaks) must be ≥2 than the peak intensities obtained from control (light peaks) to be considered quantitatively significant. Out of 44 specific HIF-1α-interacting partners, the histone demethylase jumonji domain containing protein 2C (JMJD2C) encoded by KDM4C gene, which has been previously reported to be HIF-1 target gene, was further investigated. Experimental validation by JMJD2C knockdown has proven the significance of JMJD2C in breast tumor growth and metastasis to the lungs of mice.

SILAC-based approach has been applied in proteomic profiling of secretomes. Proteins secreted from cells will be labeled with "heavy" amino acids, which can then be separated from contamination of "light" amino acids of proteins from fetal bovine serum in the cell culture media (106). As noted previously that metabolic labeling permits mixing of different samples prior to quantitation, SILAC is thus suitable for secretome research as any errors caused by extensive sample processing in isolation and enrichment steps to obtain secretome will be eliminated. In 2015, a study conducted by Erler et al. serves as a good experimental model for hypoxia cancer secretome quantitative analysis (Quadrupole-Orbitrap LC-MS/MS system) (65). This selective

sorting procedures of secretome with sophisticated MS instrumentation has provided accurate global identification and quantitation of secreted proteins that are uniquely release under different breast cancer cell types and hypoxic conditions (*i.e.* MDA-MB-231 parental 21% and 1% O_2, MDA-MB-231 Bone-Tropic 21 and 1% O_2). In this study, lysyl oxidase (LOX) was identified as one of most highly upregulated (>2.25-fold) secreted proteins associated with osteotropism in hypoxia-induced secretome release from MDA-MB-231 Bone Tropic breast cancer cell lines (107). Subsequence functional analysis *in vivo* successfully identified secreted LOX as a novel hypoxia mediator that initiates the formation of pre-metastatic bone lesion which eventually allows circulating tumor cells to colonize the bone (65).

Application of Pulsed-SILAC Labeling and LC–MS/MS for Quantitative Analysis of the Dynamic State of Hypoxia-Associated Proteins Turnover

The conventional SILAC labeling provides the knowledge of a steady-state amount of protein in the cell. Nonetheless, it is important to emphasize that the concentration of protein found does not always correlate with rate of protein synthesis, but rather the result of the balance between protein synthesis and degradation (108). For this reason, understanding dynamic proteome relationship is important when interpreting proteins abundance data from proteomic experiments to understand how cells regulate their proteomes to execute numerous cellular responses to particular stimuli.

The emergent approach to identify *de novo* synthesis of proteins and the measurement of their half-life by adapting the classic pulse-chase experiment to SILAC workflows (pSILAC) has enabled researchers to gain insights into the dynamics of protein expression on a proteome-wide scale. As a modified version of the conventional SILAC approach, instead of mixing normal light and heavy peptide samples after isotopes incorporation, pSILAC involves the switching of amino acid in the media from normal light to heavy (or *vice versa*) at a certain time point prior to perturbation such as hypoxia stress. At this time point on, proteins translating in cells are pulse-labeled and all newly synthesized proteins are distinguishable from the preexisting proteome. Then, the relative abundance of light and heavy peptides overtime can be assessed to quantify the actively translating proteins induced by the perturbation.

It was not until recently that quantitative pSILAC was first applied to study proteome in hypoxia cancer model. In 2016, a study was performed on U87MG human glioblastoma and 786-O human renal clear cell carcinoma cell lines to investigate the effect of oxygen tension on translation efficiency. The results revealed that oxygen perturbation can cause widespread changes in protein output, with minimal changes on mRNA levels. Notably, this study reported that HIF target genes show no alteration upon oxygen deprivation in contrast to an increase in translation level (109).

Later on, a global analysis of newly translated proteins between normoxic and hypoxic (24 h) A431 epidermoid carcinoma cells revealed an intriguing finding that only 5% of total proteins were upregulated, and more than 60% were translationally suppressed in hypoxia condition (66). In concordance with the previously mentioned study by Bousquet et al. [in 2015 (104) and 2018 (61)] using the traditional SILAC method, key modulators of aerobic glycolysis [*i.e.* glucose transporter 1 (GLUT1) and hexokinase 2 (HK2)] were found *de novo* synthesized for up to 1.8-fold, and that protein synthesis of key TCA cycle enzymes were suppressed. This data confirmed the metabolic switch from oxidative phosphorylation to glycolysis. However, in this latter pSILAC study, some glycolytic enzymes including HK1 and lactate dehydrogenase A that were previously found upregulated are suppressed (66). The possible explanation to this phenomenon is a slow rate of protein degradation, which resulted in the overall increase in this set of glycolysis proteins in hypoxic cells (110). Moreover, information from pSILAC proteomic analysis reveal for the first time that PHF14 is an epigenetic modulator that plays a key role in cell growth cessation, a classical cellular response in cells under hypoxia (66).

In a pancreatic cancer model, pSILAC and high resolution MS were used to identify the dynamic expression of hypoxia-induced mediators. In this study, the elevated expression of proteins that play a role in immune suppression, angiogenesis, metabolic activity, and metastasis were found. The increase in this set of proteins is associated with poor patient survival in various stages of pancreatic cancer (67). Interestingly, one of the highly inducible proteins upon hypoxia stress, ERO1α, was demonstrated by genetic deletion of ERO1α gene in pancreatic cancer cell lines and mouse xenograft model to slow down growth rate of tumor cells, indicating the significance of ERO1α in cancer progression *in vivo* and *in vitro* (67).

It is known that cancer-associated fibroblasts (CAFs) play an important role in the pathogenesis of cancer, but the response of CAFs upon oxygen restriction remains unclear. In 2019, Kugeratski et al. performed conventional SILAC-MS to characterize both proteome and secretome of hypoxic CAFs. In this study, the function of NCBP2-AS2 was highlighted as its level increased significantly in hypoxia condition. Subsequent analysis using a pSILAC-based workflow confirmed that hypoxia induces NCBP2-AS2 posttranscriptionally by enhancing its translation. Functional analysis of NCBP2-AS2 confirmed the pro-angiogenic and pro-migratory function of CAFs under hypoxic condition (68).

Chemical Labeling Coupled With MS

While the original metabolic labeling is not applicable in samples that are not metabolically active (*i.e.* biofluids and biopsies material), the introduction of isotope labeling to proteins or peptides using chemical reactions such as isotopic isotope-coded affinity tag (ICAT) (111) and isotope-coded protein labels (ICPL) (112) or isobaric tag for relation and absolute quantification (iTRAQ) (113) and tandem mass tags (TMT) (114) labeling can be done on any proteome sample (115). Moreover, chemical labeling has been reported to reduce sample processing time (116). However, major limitations include the requirement of specific proteins (*i.e.* cysteine or lysine) to quantify and chemical

labeling is performed at the late stage of sample preparation, which therefore increases the chance of errors being introduced in the multi-steps proteomic sample preparation (115). Using the chemical labeling approach, a number of studies have successfully uncovered potential cancer biomarkers as well as key players in tumorigenicity and resistance to cancer therapy.

Isobaric tag-based high-throughput quantitative proteomics is based on the covalent conjugation of the stable isotope to the N-terminus and side chain amines of the peptides. The individual samples labeled with different isobaric tags are often pooled and subjected to fractionation by liquid chromatography prior to LC-MS/MS analysis. Despite the superior quantitation accuracy of SILAC, isobaric tag methods outperforms the method of SILAC in terms of sensitivity with a capacity to detect reporter ions in the low *m/z* region and the ability to be multiplexed to up to 8 (for iTRAQ) and 16 (for TMT) samples being processed simultaneously.

Application of iTRAQ for Quantitative Analysis of Multiplexed Hypoxia Samples

The first study that incorporated the use of iTRAQ for relative quantification in hypoxia and cancer model was in 2010 (117). Park and colleagues performed iTRAQ labeling and LC-MS/MS on the secretome of hypoxic A431 squamous carcinoma cells to elucidate the tumorigenic mechanisms (117). Isobaric tags of varying mass were labeled on cells under normoxic, hypoxic 48 h, hypoxic 72 h, and 48 h hypoxic followed by 24 h reoxygenation conditions. The levels of protein expression were then quantitated using relative peak intensities of the tags. Using this approach, the results revealed significant alteration in secretome including soluble proteins and exosomes/extracellular vesicles of A431 cells under hypoxia condition which were linked to their metastatic and angiogenic potential. Notably, there was a decline in the secretion of extracellular matrix protein components that involved in focal adhesion, antiangiogenic factors, and regulator of intercellular adhesion; and an increase in potent pro-metastatic factors. This data supported the observed phenomenon that under hypoxia condition, the tumor cells exhibited reduced adhesion and enhanced invasiveness as determined by adhesion and chemoinvasion and induced angiogenesis by chorioallantoic membrane (CAM) assay.

With the same A431 cells, the differential protein expression after hypoxia and reoxygenation treatment were investigated using iTRAQ quantitative method (2013) (69). More than 4,000 proteins were identified, 1200 of which were modulated >1.2 fold. Several potential novel targets for tumor therapy were identified. This includes proteins that are involved in glycolysis pathway, STAT1 pathway, glycoprotein synthesis, and integrin (69, 70). Notably, the upregulation of non-homologous end-joining pathway which plays role in DNA repair in hypoxic A431 cells were quantified. iTRAQ labeling has also been used to identify radiotherapy-resistant protein targets. Quantification of the change in proteome of irradiated hypoxic A431 cells by low-dose γ irradiation showed a significant upregulation of several calcium binding proteins (71). Subsequent knockdown using shRNA of genes encoding these proteins resulted in hypoxic A431 cells which were resistant to radiation. These findings have, for the first time, identified radiotherapy-resistant protein targets for cancer patients undergoing radiotherapy.

The robust and sensitive iTRAQ quantification has quickly become a popular method to quantify sub-cellular fractionation proteins (118). More recently, iTRAQ labeling has been used to compare the profile of chromatome, chromatin-associated proteins, in normoxic, hypoxic and re-oxygenated A431 cells (19). Chromatin is a complex structure containing both DNA and proteins which function to regulate cellular processes that requires access to DNA (*i.e.* DNA replication, DNA transcription, DNA-repair). Although it is known that the transcriptional change during hypoxia is mainly regulated by the HIF, emerging evidence has shown that gene transcriptional changes exerted by other non-dependent chromatome-induced epigenetic changes also play a key part of the hypoxia response (25, 119, 120). Therefore, identification of chromatin-associated proteins is essential in understanding the cellular response to hypoxia environment. The iTRAQ method for quantitation of differentially regulated chromatome has shown that 819 proteins changed their chromatin association topology under hypoxic conditions. Functional investigation of the chromatin organizer protein HP1BP3 has revealed its role in regulating chromatin condensation upon hypoxia treatment, which can mediate tumor progression and acquire therapy-resistant traits (19). Furthermore, iTRAQ-based proteomics to quantify chromatin-associated proteins during the interphase of cell cycle has again identified HP1BP3 to be a key player that control the cells ability to proliferate (72).

Application of TMT for Quantitative Analysis of Hypoxic Cancer Cells Proteome/Secretome

In 2017, TMT was utilized along with label-free LC-MS/MS to analyze the hypoxic and normoxic proteome of U87 glioblastoma cells (73). Zhang et al. adapted the SILAC technique to TMT and label-free proteomics in replicating cells by first incorporating light arginine to the proteins instead of the heavy isotope-containing arginine for 5 days, followed by TMT stable isotope labeling. This way, not only did the expression of newly synthesized proteins can be quantified, but the cell proliferation rates can also be monitored. The proteins quantified by TMT proteomics revealed that hypoxia upregulates glucose transport and glycolysis pathways. From the proteomic data, glucose transporter (GLUT1), DUSP4/MKP2, and RelA proteins which are involved in inflammation, enzymes for glycolytic pathway, and proteins involved in cells transition from an epithelial phenotype to a type III EMT mesenchymal phenotype were elevated. In addition, a novel finding was found that under hypoxic condition, the vitamin B_{12} transporter protein TCN2 is significantly downregulated, resulting in cells' growth arrested and also plays an important role in directing cancer cell transformation toward the highly aggressive mesenchymal/CSC stage.

TMT 6-plex tagging of three independent sets of normoxic and hypoxic exosomal protein released from glioblastoma

(GBM) cells was performed by Kora et al. in 2018 to elucidate the mechanisms of cancer progression (75). The results revealed that hypoxia condition significantly induced the upregulation of various exosomal proteins that were reported to be involved in tumor progression, metastasis and angiogenesis. In another study, TMT-based analysis was performed on exosome isolated from B16-F0 mouse melanoma cell culture supernatants of normoxic and hypoxic conditions (74). The result showed an enrichment of proteins that are associated with immune responses. Concomitant with MS result, the isolated exosomes has been demonstrated to influence the macrophage M2-like polarization and promote oxidative phosphorylation in bone marrow-derived macrophages *via* transfer for let-7a miRNA, which resulted in the inhibition of the insulin–Akt–mTOR pathway. These findings provided multiple targets including the signaling pathway and proteins that contribute to cancer metastasis and that hypoxia-induced exosomal release from tumor cells may be a source of potential biomarkers for cancer diagnosis (121).

Multiple Reaction Monitoring

MRM is a sensitive and highly specific MS technique that has been extensively used for the quantification of small molecules (87, 115), for example, the metabolite profiling of clinical samples (122). In typical MRM experiments, the target precursor ions are fragmented and only the pre-selected fragment ions of interest are being monitored (123). Likewise, the same principle of quantification is applicable to proteomic applications. MRM is one of the most commonly used technique in targeted proteomic assays. This specific selective quantification of peptide of interest provide researchers with a tremendous advantage especially in verifying proteomic markers when specific antibodies for immune assay are not available.

The use of MRM has also been applied for hypoxia and cancer analysis. As discussed previously, proteome change identified by iTRAQ labeling coupled with LC-MS/MS of A431 epithelial carcinoma cells induced by hypoxia and reoxygenation revealed a set of proteins perturbed by hypoxia and not in normoxia (69). This set of unique protein expression was subsequently confirmed using the LC-MS/MS-MRM quantification method, in which most of the proteins are verified. Notably, upregulation of key proteins in the nonhomologous end-joining pathway, Ku70/Ku80 dimer, was reported to be much higher than the quantification result obtained from iTRAQ (ratio 2.05 vs. 1.1–1.2, respectively). In this specific experimental setting, this finding demonstrated the superior sensitivity of MRM method compared to iTRAQ labeling.

LABEL-FREE MS-BASED STRATEGIES

With the advancement in MS, the use of label-free quantitation has increased over the past few years. Numerous studies demonstrated the success in using label-free method for relative protein quantitation to investigate how hypoxia drives cancer progression (73, 124–127). Label-free MS-based protein quantification is commonly used in proteomic analysis. The major advantage in using label-free quantitation is that it requires no labeling of the sample and applicable to all types of proteomic samples. Two of the most widely performed methods are spectral counting and peptide peak intensity measurement. Spectral counting is more suitable for proteins with high number and high abundance in the samples. Only sufficient amounts of peptides to trigger MS/MS spectra are quantifiable. Therefore, spectral counting is less reliable for low mass and low abundant proteins (128). On the other hand, quantitation of proteins by peptide peak intensity compares the total amount of peptide ion intensities using extracted ion chromatogram in a specific retention time computed and showed as the area under the curve. However, it is reported that up to 40% of data variation at the peptide level can occur since the coverage of common ions are strongly dependent on the amount of sample loaded, column condition, and calibration. Thus, achieving reproducibility is difficult between different sample runs (129).

CONCLUDING REMARKS

The field of proteomics is based on the analysis of large number of proteins simultaneously. The use of MS is a powerful technology that facilitates the success of quantitative and differential proteome analysis. The number of publications on hypoxia and cancer is escalating with increasing in knowledge in this topic. Choosing the right technique depends on research questions being addressed as one methodology may be better than another for one's particular experiment. It is necessary for scientists to be able to keep up with the rapid advancement in new analytical methods and instrumentation as well as understanding the principles, strengths and limitations of the methods.

However, there is room for improvement in MS-based proteomics. The algorithms of peptide identification and bioinformatics computational big data analysis are two areas that can be improved. Since peptides obtained from the experimental MS/MS will be assigned to theoretical peptide sequences in a given protein database, implementation of appropriate algorithms would reduce the number of incorrect peptide matches and, therefore providing protein lists of high confidence. Cancer proteome comprises extremely complex proteoforms including mutated proteins, cancer specific PTMs, hypoxia induced proteins and PTMs etc. Bottom up approach is the commonly used method in cancer proteomic research by digesting proteins to peptides using specific protease because peptides are easier to handle and separate in LC-MS/MS analysis. However, the main disadvantage of the peptide centric bottom up approach is the protein inference problem, *i.e.* the detected peptide can be presented in multiple different proteins, leading to ambiguities in determining the proteins and their biological functions. Top down mass spectrometry which measures and sequences individual intact proteins directly without digestion can precisely characterize proteoforms including mutations and post translational modifications. Thus top down approach is an essential method for studying hypoxia cancer biology and should

be further developed to precisely determine proteoforms in a cancer sample for accurate interpretation of biological functions.

Despite tremendous progress of analytical hardware, tools for biological and functional annotation of data are still in its infancy. Typically, data obtained from MS are highly complex. It needs to be emphasized that choosing software applications for downstream analysis is as equally important as sample preparation and MS analysis, considering that it determines the meaningfulness of the result and influences how one interpret the data. This is set to improve with the continuous incorporation of new technologies.

Another major point that is equally important in studying hypoxia and cancer is the experimental design to create the hypoxic environment for experimentation. It has been found that the biological consequences largely depends on the time of cells exposure to hypoxia (130). Generally with similar observing trends, cancer cells under acute hypoxia *in vitro* ranging between a few minutes to 72 h have been shown to exhibit apoptotic and metabolic adaptation leading to tumor cell survival and progression (131, 132), while cells under chronic hypoxia ranging from a few hours to many weeks have demonstrated the high frequency of DNA breaks leading to genomic instability and mutagenesis (133), both of which have been shown to increase radio-resistance in cancer (134). Nonetheless, it needs to be emphasized that these observed responses are not exclusive to acute or chronic hypoxia, and that inconsistent findings have been reported in the past (130). For this reason, standardization of the time of hypoxia exposure and the identification of the proteins underlining specific time points of treatment is important.

We have discussed the technological advancements, future directions as well as challenges in the application of mass spectrometry-based proteomics in the context of cancer and hypoxia research. With continuous improvement and newer technologies being introduced, it is very likely that there will be increased utilization of proteomics for diagnostic biomarker and therapeutic discovery in cancer.

AUTHOR CONTRIBUTIONS

AV drafted the manuscript. JL, KY, WC, and SS revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work is in part supported by the Singapore Ministry of Education (MOE2016-T2-2-018 and MOE2018-T1-001-078), and Singapore National Medical Research Council (NMRC/OFIRG/0003/2016).

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* (2019) 69(1):7–34. doi: 10.3322/caac.21551
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
3. Patel C, Brotherton JM, Pillsbury A, Jayasinghe S, Donovan B, Macartney K, et al. The impact of 10 years of human papillomavirus (HPV) vaccination in Australia: what additional disease burden will a nonavalent vaccine prevent? *Euro Surveill* (2018) 23(41):30–3. doi: 10.2807/1560-7917.ES.2018.23.41.1700737
4. Smith ER, George SH, Kobetz E, Xu XX. New biological research and understanding of Papanicolaou's test. *Diagn Cytopathol* (2018) 46(6):507–15. doi: 10.1002/dc.23941
5. Simms KT, Steinberg J, Caruana M, Smith MA, Lew JB, Soerjomataram I, et al. Impact of scaled up human papillomavirus vaccination and cervical screening and the potential for global elimination of cervical cancer in 181 countries, 2020-99: a modelling study. *Lancet Oncol* (2019) 20(3):394–407. doi: 10.1016/s1470-2045(18)30836-2
6. Gambardella V, Tarazona N, Cejalvo JM, Lombardi P, Huerta M, Rosello S, et al. Personalized Medicine: Recent Progress in Cancer Therapy. *Cancers* (2020) 12(4):1–9. doi: 10.3390/cancers12041009
7. Greaves M, Maley CC. Clonal evolution in cancer. *Nature* (2012) 481 (7381):306–13. doi: 10.1038/nature10762
8. Derbal Y. Perspective on the dynamics of cancer. *Theor Biol Med Model* (2017) 14(1):18. doi: 10.1186/s12976-017-0066-5
9. Robertson-Tessi M, Anderson AR. Big Bang and context-driven collapse. *Nat Genet* (2015) 47(3):196–7. doi: 10.1038/ng.3231
10. Navin NE. The first five years of single-cell cancer genomics and beyond. *Genome Res* (2015) 25(10):1499–507. doi: 10.1101/gr.191098.115
11. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet* (2016) 17(3):175–88. doi: 10.1038/nrg.2015.16
12. Roerink SF, Sasaki N, Lee-Six H, Young MD, Alexandrov LB, Behjati S, et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* (2018) 556(7702):457–62. doi: 10.1038/s41586-018-0024-3
13. Yu L, Zhao H, Meng L, Zhang C. Application of Single Cell Sequencing in Cancer. *Adv Exp Med Biol* (2018) 1068:135–48. doi: 10.1007/978-981-13-0502-3_11
14. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* (2016) 534(7605):55–62. doi: 10.1038/nature18003
15. Petrova V, Annicchiarico-Petruzzelli M, Melino G, Amelio I. The hypoxic tumour microenvironment. *Oncogenesis* (2018) 7(1):10. doi: 10.1038/s41389-017-0011-9
16. Ruan K, Song G, Ouyang G. Role of hypoxia in the hallmarks of human cancer. *J Cell Biochem* (2009) 107(6):1053–62. doi: 10.1002/jcb.22214
17. Muz B, de la Puente P, Azab F, Azab AK. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia (Auckl)* (2015) 3:83–92. doi: 10.2147/HP.S93413
18. Gray LH, Conger AD, Ebert M, Hornsey S, Scott OC. The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br J Radiol* (1953) 26(312):638–48. doi: 10.1259/0007-1285-26-312-638
19. Dutta B, Yan R, Lim SK, Tam JP, Sze SK. Quantitative profiling of chromatome dynamics reveals a novel role for HP1BP3 in hypoxia-induced oncogenesis. *Mol Cell Proteom* (2014) 13(12):3236–49. doi: 10.1074/mcp.M114.038232
20. Lee P, Chandel NS, Simon MC. Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. *Nat Rev Mol Cell Biol* (2020) 21:268–83. doi: 10.1038/s41580-020-0227-y
21. Jun JC, Rathore A, Younas H, Gilkes D, Polotsky VY. Hypoxia-Inducible Factors and Cancer. *Curr Sleep Med Rep* (2017) 3(1):1–10. doi: 10.1007/s40675-017-0062-7
22. Vaupel P. The role of hypoxia-induced factors in tumor progression. *Oncologist* (2004) 9(Suppl 5):10–7. doi: 10.1634/theoncologist.9-90005-10

23. Semenza GL. Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene* (2010) 29(5):625–34. doi: 10.1038/onc.2009.441
24. Semenza GL, Agani F, Booth G, Forsythe J, Iyer N, Jiang BH, et al. Structural and functional analysis of hypoxia-inducible factor 1. *Kidney Int* (1997) 51 (2):553–5. doi: 10.1038/ki.1997.77
25. Melvin A, Rocha S. Chromatin as an oxygen sensor and active player in the hypoxia response. *Cell Signal* (2012) 24(1):35–43. doi: 10.1016/j.cellsig.2011.08.019
26. Agani F, Jiang BH. Oxygen-independent regulation of HIF-1: novel involvement of PI3K/AKT/mTOR pathway in cancer. *Curr Cancer Drug Targ* (2013) 13(3):245–51. doi: 10.2174/1568009611313030003
27. Courtnay R, Ngo DC, Malik N, Ververis K, Tortorella SM, Karagiannis TC. Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. *Mol Biol Rep* (2015) 42(4):841–51. doi: 10.1007/s11033-015-3858-x
28. D'Ignazio L, Rocha S. Hypoxia Induced NF-kappaB. *Cells* (2016) 5(1):1–5. doi: 10.3390/cells5010010
29. Minet E, Arnould T, Michel G, Roland I, Mottet D, Raes M, et al. ERK activation upon hypoxia: involvement in HIF-1 activation. *FEBS Lett* (2000) 468(1):53–8. doi: 10.1016/S0014-5793(00)01181-9
30. Krock BL, Skuli N, Simon MC. Hypoxia-induced angiogenesis: good and evil. *Genes Cancer* (2011) 2(12):1117–33. doi: 10.1177/1947601911423654
31. Carmeliet P. Angiogenesis in life, disease and medicine. *Nature* (2005) 438 (7070):932–6. doi: 10.1038/nature04478
32. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. *Nature* (2011) 473(7347):298–307. doi: 10.1038/nature10144
33. Ouyang L, Shi Z, Zhao S, Wang FT, Zhou TT, Liu B, et al. Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. *Cell Prolif* (2012) 45(6):487–98. doi: 10.1111/j.1365-2184.2012.00845.x
34. Qiu Y, Li P, Ji C. Cell Death Conversion under Hypoxic Condition in Tumor Development and Therapy. *Int J Mol Sci* (2015) 16(10):25536–51. doi: 10.3390/ijms161025536
35. Akay C, Thomas C,3, Gazitt Y. Arsenic trioxide and paclitaxel induce apoptosis by different mechanisms. *Cell Cycle* (2004) 3(3):324–34. doi: 10.4161/cc.3.3.657
36. Hassan M, Watari H, AbuAlmaaty A, Ohba Y, Sakuragi N. Apoptosis and molecular targeting therapy in cancer. *BioMed Res Int* (2014) 2014:150845. doi: 10.1155/2014/150845
37. Wouters A, Pauwels B, Lardon F, Vermorken JB. Review: implications of in vitro research on the effect of radiotherapy and chemotherapy under hypoxic conditions. *Oncologist* (2007) 12(6):690–712. doi: 10.1634/theoncologist.12-6-690
38. Teicher BA, Lazo JS, Sartorelli AC. Classification of antineoplastic agents by their selective toxicities toward oxygenated and hypoxic tumor cells. *Cancer Res* (1981) 41(1):73–81.
39. Teicher BA, Holden SA, al-Achi A, Herman TS. Classification of antineoplastic treatments by their differential toxicity toward putative oxygenated and hypoxic tumor subpopulations in vivo in the FSaIIC murine fibrosarcoma. *Cancer Res* (1990) 50(11):3339–44.
40. Strese S, Fryknas M, Larsson R, Gullbo J. Effects of hypoxia on human cancer cell line chemosensitivity. *BMC Cancer* (2013) 13:331. doi: 10.1186/1471-2407-13-331
41. Brown JM. Exploiting the hypoxic cancer cell: mechanisms and therapeutic strategies. *Mol Med Today* (2000) 6(4):157–62. doi: 10.1016/S1357-4310(00)01677-4
42. Moyer MW. Targeting hypoxia brings breath of fresh air to cancer therapy. *Nat Med* (2012) 18(5):636–7. doi: 10.1038/nm0512-636b
43. Wigerup C, Pahlman S, Bexell D. Therapeutic targeting of hypoxia and hypoxia-inducible factors in cancer. *Pharmacol Ther* (2016) 164:152–69. doi: 10.1016/j.pharmthera.2016.04.009
44. Hudson CC, Liu M, Chiang GG, Otterness DM, Loomis DC, Kaper F, et al. Regulation of hypoxia-inducible factor 1alpha expression and function by the mammalian target of rapamycin. *Mol Cell Biol* (2002) 22(20):7004–14. doi: 10.1128/mcb.22.20.7004-7014.2002
45. Terzuoli E, Puppo M, Rapisarda A, Uranchimeg B, Cao L, Burger AM, et al. Aminoflavone, a ligand of the aryl hydrocarbon receptor, inhibits HIF-1alpha expression in an AhR-independent fashion. *Cancer Res* (2010) 70 (17):6837–48. doi: 10.1158/0008-5472.CAN-10-1075
46. Lee K, Zhang H, Qian DZ, Rey S, Liu JO, Semenza GL. Acriflavine inhibits HIF-1 dimerization, tumor growth, and vascularization. *Proc Natl Acad Sci U S A* (2009) 106(42):17910–5. doi: 10.1073/pnas.0909353106
47. Kaluz S, Kaluzova M, Stanbridge EJ. Proteasomal inhibition attenuates transcriptional activity of hypoxia-inducible factor 1 (HIF-1) via specific effect on the HIF-1alpha C-terminal activation domain. *Mol Cell Biol* (2006) 26(15):5895–907. doi: 10.1128/MCB.00552-06
48. Denny WA. The role of hypoxia-activated prodrugs in cancer therapy. *Lancet Oncol* (2000) 1(1):25–9. doi: 10.1016/S1470-2045(00)00006-1
49. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science* (1989) 246 (4926):64–71. doi: 10.1126/science.2675315
50. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* (1988) 60 (20):2299–301. doi: 10.1021/ac00171a028
51. Yates JR,3. Mass spectral analysis in proteomics. *Annu Rev Biophys Biomol Struct* (2004) 33:297–316. doi: 10.1146/annurev.biophys.33.111502.082538
52. Senko MW, Speir JP, McLafferty FW. Collisional activation of large multiply charged ions using Fourier transform mass spectrometry. *Anal Chem* (1994) 66(18):2801–8. doi: 10.1021/ac00090a003
53. Shukla AK, Futrell JH. Tandem mass spectrometry: dissociation of ions by collisional activation. *J Mass Spectrom* (2000) 35(9):1069–90. doi: 10.1002/1096-9888(200009)35:9<1069::AID-JMS54>3.0.CO;2-C
54. Fung YM, Adams CM, Zubarev RA. Electron ionization dissociation of singly and multiply charged peptides. *J Am Chem Soc* (2009) 131(29):9977–85. doi: 10.1021/ja8087407
55. Zubarev RA, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, Lewis MA, et al. Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem* (2000) 72(3):563–73. doi: 10.1021/ac990811p
56. Sze SK, Ge Y, Oh H, McLafferty FW. Plasma electron capture dissociation for the characterization of large proteins by top down mass spectrometry. *Anal Chem* (2003) 75(7):1599–603. doi: 10.1021/ac020446t
57. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* (2004) 101(26):9528–33. doi: 10.1073/pnas.0402700101
58. Li X, Wang W, Chen J. Recent progress in mass spectrometry proteomics for biomedical research. *Sci China Life Sci* (2017) 60(10):1093–113. doi: 10.1007/s11427-017-9175-2
59. O'Neill JR. An Overview of Mass Spectrometry-Based Methods for Functional Proteomics. *Methods Mol Biol* (2019) 1871:179–96. doi: 10.1007/978-1-4939-8814-3_13
60. Domon B, Aebersold R. Mass spectrometry and protein analysis. *Science* (2006) 312(5771):212–7. doi: 10.1126/science.1124619
61. Bousquet PA, Sandvik JA, Jeppesen Edin NF, Krengel U. Hypothesis: Hypoxia induces de novo synthesis of NeuGc gangliosides in humans through CMAH domain substitute. *Biochem Biophys Res Commun* (2018) 495(1):1562–6. doi: 10.1016/j.bbrc.2017.11.183
62. Yamada K, Ono M, Perkins ND, Rocha S, Lamond AI. Identification and functional characterization of FMN2, a regulator of the cyclin-dependent kinase inhibitor p21. *Mol Cell* (2013) 49(5):922–33. doi: 10.1016/j.molcel.2012.12.023
63. Greenhough A, Bagley C, Heesom KJ, Gurevich DB, Gay D, Bond M, et al. Cancer cell adaptation to hypoxia involves a HIF-GPRC5A-YAP axis. *EMBO Mol Med* (2018) 10(11):1–13. doi: 10.15252/emmm.201708699
64. Luo W, Chang R, Zhong J, Pandey A, Semenza GL. Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression. *Proc Natl Acad Sci U S A* (2012) 109(49):E3367–76. doi: 10.1073/pnas.1217394109
65. Cox TR, Rumney RMH, Schoof EM, Perryman L, Hoye AM, Agrawal A, et al. The hypoxic cancer secretome induces pre-metastatic bone lesions through lysyl oxidase. *Nature* (2015) 522(7554):106–10. doi: 10.1038/nature14492
66. Park JE, Tse SW, Xue G, Assisi C, Maqueda AS, Ramon GPX, et al. Pulsed SILAC-based proteomic analysis unveils hypoxia- and serum starvation-induced de novo protein synthesis with PHD finger protein 14 (PHF14) as a hypoxia sensitive epigenetic regulator in cell cycle progression. *Oncotarget* (2019) 10(22):2136–50. doi: 10.18632/oncotarget.26669

67. Gupta N, Park JE, Tse W, Low JK, Kon OL, McCarthy N, et al. ERO1alpha promotes hypoxic tumor progression and is associated with poor prognosis in pancreatic cancer. *Oncotarget* (2019) 10(57):5970–82. doi: 10.18632/oncotarget.27235
68. Kugeratski FG, Atkinson SJ, Neilson LJ, Lilla S, Knight JRP, Serneels J, et al. Hypoxic cancer-associated fibroblasts increase NCBP2-AS2/HIAR to promote endothelial sprouting through enhanced VEGF signaling. *Sci Signal* (2019) 12(567):1–33. doi: 10.1126/scisignal.aan8247
69. Ren Y, Hao P, Dutta B, Cheow ES, Sim KH, Gan CS, et al. Hypoxia modulates A431 cellular pathways association to tumor radioresistance and enhanced migration revealed by comprehensive proteomic and functional studies. *Mol Cell Proteomics* (2013) 12(2):485–98. doi: 10.1074/mcp.M112.018325
70. Ren Y, Hao P, Law SK, Sze SK. Hypoxia-induced changes to integrin alpha 3 glycosylation facilitate invasion in epidermoid carcinoma cell line A431. *Mol Cell Proteom* (2014) 13(11):3126–37. doi: 10.1074/mcp.M114.038505
71. Ren Y, Yeoh KW, Hao P, Kon OL, Sze SK. Irradiation of Epithelial Carcinoma Cells Upregulates Calcium-Binding Proteins That Promote Survival under Hypoxic Conditions. *J Proteome Res* (2016) 15(12):4258–64. doi: 10.1021/acs.jproteome.6b00340
72. Dutta B, Ren Y, Hao P, Sim KH, Cheow E, Adav S, et al. Profiling of the Chromatin-associated Proteome Identifies HP1BP3 as a Novel Regulator of Cell Cycle Progression. *Mol Cell Proteomics* (2014) 13(9):2183–97. doi: 10.1074/mcp.M113.034975
73. Zhang K, Xu P, Sowers JL, Machuca DF, Mirfattah B, Herring J, et al. Proteome Analysis of Hypoxic Glioblastoma Cells Reveals Sequential Metabolic Adaptation of One-Carbon Metabolic Pathways. *Mol Cell Proteomics* (2017) 16(11):1906–21. doi: 10.1074/mcp.RA117.000154
74. Park JE, Dutta B, Tse SW, Gupta N, Tan CF, Low JK, et al. Hypoxia-induced tumor exosomes promote M2-like macrophage polarization of infiltrating myeloid cells and microRNA-mediated metabolic shift. *Oncogene* (2019) 38(26):5158–73. doi: 10.1038/s41388-019-0782-x
75. Kore RA, Edmondson JL, Jenkins SV, Jamshidi-Parsian A, Dings RPM, Reyna NS, et al. Hypoxia-derived exosomes induce putative altered pathways in biosynthesis and ion regulatory channels in glioblastoma cells. *Biochem Biophys Rep* (2018) 14:104–13. doi: 10.1016/j.bbrep.2018.03.008
76. Gao Z, Luo G, Ni B. Progress in mass spectrometry-based proteomic research of tumor hypoxia (Review). *Oncol Rep* (2017) 38(2):676–84. doi: 10.3892/or.2017.5748
77. Cui Z, Dewey S, Gomes AV. Cardioproteomics: advancing the discovery of signaling mechanisms involved in cardiovascular diseases. *Am J Cardiovasc Dis* (2011) 1(3):274–92.
78. Sze SK. Artificially intelligent proteomics improves cardiovascular risk assessment. *EBioMedicine* (2019) 40:23–4. doi: 10.1016/j.ebiom.2019.01.014
79. Gerszten RE, Asnani A, Carr SA. Status and prospects for discovery and verification of new biomarkers of cardiovascular disease by proteomics. *Circ Res* (2011) 109(4):463–74. doi: 10.1161/CIRCRESAHA.110.225003
80. Le QT, Courter D. Clinical biomarkers for hypoxia targeting. *Cancer Metastasis Rev* (2008) 27(3):351–62. doi: 10.1007/s10555-008-9144-9
81. Belczacka I, Latosinska A, Metzger J, Marx D, Vlahou A, Mischak H, et al. Proteomics biomarkers for solid tumors: Current status and future prospects. *Mass Spectrom Rev* (2019) 38(1):49–78. doi: 10.1002/mas.21572
82. Lubec G, Afjehi-Sadat L. Limitations and pitfalls in protein identification by mass spectrometry. *Chem Rev* (2007) 107(8):3568–84. doi: 10.1021/cr068213f
83. Buchberger AR, DeLaney K, Johnson J, Li L. Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* (2018) 90(1):240–65. doi: 10.1021/acs.analchem.7b04733
84. Borras E, Sabido E. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. *Proteomics* (2017) 17(17-18):1–8. doi: 10.1002/pmic.201700180
85. Bronstrup M. Absolute quantification strategies in proteomics based on mass spectrometry. *Expert Rev Proteomics* (2004) 1(4):503–12. doi: 10.1586/14789450.1.4.503
86. Gevaert K, Impens F, Ghesquiere B, Van Damme P, Lambrechts A, Vandekerckhove J. Stable isotopic labeling in proteomics. *Proteomics* (2008) 8(23-24):4873–85. doi: 10.1002/pmic.200800421
87. Wasinger VC, Zeng M, Yau Y. Current status and advances in quantitative proteomic mass spectrometry. *Int J Proteom* (2013) 2013:180605. doi: 10.1155/2013/180605
88. Han X, Aslanian A, Yates JR,3. Mass spectrometry for proteomics. *Curr Opin Chem Biol* (2008) 12(5):483–90. doi: 10.1016/j.cbpa.2008.07.024
89. Ranish JA, Yi EC, Leslie DM, Purvine SO, Goodlett DR, Eng J, et al. The study of macromolecular complexes by quantitative proteomics. *Nat Genet* (2003) 33(3):349–55. doi: 10.1038/ng1101
90. Martinez-Fabregas J, Rubio S, Diaz-Quintana A, Diaz-Moreno I, De la Rosa MA. Proteomic tools for the analysis of transient interactions between metalloproteins. *FEBS J* (2011) 278(9):1401–10. doi: 10.1111/j.1742-4658.2011.08061.x
91. Mermelekas G, Vlahou A, Zoidakis J. SRM/MRM targeted proteomics as a tool for biomarker validation and absolute quantification in human urine. *Expert Rev Mol Diagn* (2015) 15(11):1441–54. doi: 10.1586/14737159.2015.1093937
92. Schiess R, Wollscheid B, Aebersold R. Targeted proteomic strategy for clinical biomarker discovery. *Mol Oncol* (2009) 3(1):33–44. doi: 10.1016/j.molonc.2008.12.001
93. Wolf-Yadlin A, Hautaniemi S, Lauffenburger DA, White FM. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc Natl Acad Sci U S A* (2007) 104(14):5860–5. doi: 10.1073/pnas.0608638104
94. Kruger M, Kratchmarova I, Blagoev B, Tseng YH, Kahn CR, Mann M. Dissection of the insulin signaling pathway via quantitative phosphoproteomics. *Proc Natl Acad Sci U S A* (2008) 105(7):2451–6. doi: 10.1073/pnas.0711713105
95. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci U S A* (1999) 96(12):6591–6. doi: 10.1073/pnas.96.12.6591
96. Gouw JW, Tops BB, Krijgsveld J. Metabolic labeling of model organisms using heavy nitrogen (15N). *Methods Mol Biol* (2011) 753:29–42. doi: 10.1007/978-1-61779-148-2_2
97. Gouw JW, Krijgsveld J, Heck AJ. Quantitative proteomics by metabolic labeling of model organisms. *Mol Cell Proteom* (2010) 9(1):11–24. doi: 10.1074/mcp.R900001-MCP200
98. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* (2002) 1(5):376–86. doi: 10.1074/mcp.M200025-MCP200
99. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat Methods* (2010) 7(5):383–5. doi: 10.1038/nmeth.1446
100. Lin YT, Chien KY, Wu CC, Chang WY, Chu LJ, Chen MC, et al. Super-SILAC mix coupled with SIM/AIMS assays for targeted verification of phosphopeptides discovered in a large-scale phosphoproteome analysis of hepatocellular carcinoma. *J Proteomics* (2017) 157:40–51. doi: 10.1016/j.jprot.2017.02.005
101. Noberini R, Bonaldi T. A Super-SILAC Strategy for the Accurate and Multiplexed Profiling of Histone Posttranslational Modifications. *Methods Enzymol* (2017) 586:311–32. doi: 10.1016/bs.mie.2016.09.036
102. Venable JD, Wohlschlegel J, McClatchy DB, Park SK, Yates JR,3. Relative quantification of stable isotope labeled peptides using a linear ion trap-Orbitrap hybrid mass spectrometer. *Anal Chem* (2007) 79(8):3056–64. doi: 10.1021/ac062054i
103. Djidja MC, Chang J, Hadjiprocopis A, Schmich F, Sinclair J, Mrsnik M, et al. Identification of hypoxia-regulated proteins using MALDI-mass spectrometry imaging combined with quantitative proteomics. *J Proteome Res* (2014) 13(5):2297–313. doi: 10.1021/pr401056c
104. Bousquet PA, Sandvik JA, Arntzen MO, Jeppesen Edin NF, Christoffersen S, Krengel U, et al. Hypoxia Strongly Affects Mitochondrial Ribosomal Proteins and Translocases, as Shown by Quantitative Proteomics of HeLa Cells. *Int J Proteomics* (2015) 2015:678527. doi: 10.1155/2015/678527
105. Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci* (2016) 41(3):211–8. doi: 10.1016/j.tibs.2015.12.001
106. Thomas SN, Liao Z, Clark D, Chen Y, Samadani R, Mao L, et al. Exosomal Proteome Profiling: A Potential Multi-Marker Cellular Phenotyping Tool to Characterize Hypoxia-Induced Radiation Resistance in Breast Cancer. *Proteomes* (2013) 1(2):87–108. doi: 10.3390/proteomes1020087

107. Cox TR, Schoof EM, Gartland A, Erler JT, Linding R. Dataset for the proteomic inventory and quantitative analysis of the breast cancer hypoxic secretome associated with osteotropism. *Data Brief* (2015) 5:621–5. doi: 10.1016/j.dib.2015.09.039
108. Pratt JM, Petty J, Riba-Garcia I, Robertson DH, Gaskell SJ, Oliver SG, et al. Dynamics of protein turnover, a missing dimension in proteomics. *Mol Cell Proteomics* (2002) 1(8):579–91. doi: 10.1074/mcp.m200046-mcp200
109. Ho JJD, Wang M, Audas TE, Kwon D, Carlsson SK, Timpano S, et al. Systemic Reprogramming of Translation Efficiencies on Oxygen Stimulus. *Cell Rep* (2016) 14(6):1293–300. doi: 10.1016/j.celrep.2016.01.036
110. Beynon RJ. The dynamics of the proteome: strategies for measuring protein turnover on a proteome-wide scale. *Brief Funct Genom Proteomic* (2005) 3 (4):382–90. doi: 10.1093/bfgp/3.4.382
111. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* (1999) 17(10):994–9. doi: 10.1038/13690
112. Schmidt A, Kellermann J, Lottspeich F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* (2005) 5(1):4–15. doi: 10.1002/pmic.200400873
113. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* (2004) 3(12):1154–69. doi: 10.1074/mcp.M400129-MCP200
114. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* (2003) 75(8):1895–904. doi: 10.1021/ac0262560
115. Nikolov M, Schmidt C, Urlaub H. Quantitative mass spectrometry-based proteomics: an overview. *Methods Mol Biol* (2012) 893:85–100. doi: 10.1007/978-1-61779-885-6_7
116. Mirza SP, Greene AS, Olivier M. 18O labeling over a coffee break: a rapid strategy for quantitative proteomics. *J Proteome Res* (2008) 7(7):3042–8. doi: 10.1021/pr800018g
117. Park JE, Tan HS, Datta A, Lai RC, Zhang H, Meng W, et al. Hypoxic tumor cell modulates its microenvironment to enhance angiogenic and metastatic potential by secretion of proteins and exosomes. *Mol Cell Proteomics* (2010) 9(6):1085–99. doi: 10.1074/mcp.M900381-MCP200
118. Warren CM, Geenen DL, Helseth DLJr., Xu H, Solaro RJ. Sub-proteomic fractionation, iTRAQ, and OFFGEL-LC-MS/MS approaches to cardiac proteomics. *J Proteom* (2010) 73(8):1551–61. doi: 10.1016/j.jprot.2010.03.016
119. Batie M, Del Peso L, Rocha S. Hypoxia and Chromatin: A Focus on Transcriptional Repression Mechanisms. *Biomedicines* (2018) 6(2):1–19. doi: 10.3390/biomedicines6020047
120. Rocha S. Gene regulation under low oxygen: holding your breath for transcription. *Trends Biochem Sci* (2007) 32(8):389–97. doi: 10.1016/j.tibs.2007.06.005
121. Lin Y, Xu J, Lan H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. *J Hematol Oncol* (2019) 12(1):76. doi: 10.1186/s13045-019-0760-3
122. Want EJ, Cravatt BF, Siuzdak G. The expanding role of mass spectrometry in metabolite profiling and characterization. *Chembiochem* (2005) 6(11):1941–51. doi: 10.1002/cbic.200500151
123. James A, Jorgensen C. Basic design of MRM assays for peptide quantification. *Methods Mol Biol* (2010) 658:167–85. doi: 10.1007/978-1-60761-780-8_10
124. Kilic T, Valinhas ATS, Wall I, Renaud P, Carrara S. Label-free detection of hypoxia-induced extracellular vesicle secretion from MCF-7 cells. *Sci Rep* (2018) 8(1):9402. doi: 10.1038/s41598-018-27203-9
125. Nakurte I, Jekabsons K, Rembergs R, Zandberga E, Abols A, Line A, et al. Colorectal Cancer Cell Line SW480 and SW620 Released Extravascular Vesicles: Focus on Hypoxia-induced Surface Proteome Changes. *Anticancer Res* (2018) 38(11):6133–8. doi: 10.21873/anticanres.12965
126. Peixoto A, Fernandes E, Gaiteiro C, Lima L, Azevedo R, Soares J, et al. Hypoxia enhances the malignant nature of bladder cancer cells and concomitantly antagonizes protein O-glycosylation extension. *Oncotarget* (2016) 7(39):63138–57. doi: 10.18632/oncotarget.11257
127. Rodriguez J, Herrero A, Li S, Rauch N, Quintanilla A, Wynne K, et al. PHD3 Regulates p53 Protein Stability by Hydroxylating Proline 359. *Cell Rep* (2018) 24(5):1316–29. doi: 10.1016/j.celrep.2018.06.108
128. Arike L, Peil L. Spectral counting label-free proteomics. *Methods Mol Biol* (2014) 1156:213–22. doi: 10.1007/978-1-4939-0685-7_14
129. Schulze WX, Usadel B. Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* (2010) 61:491–516. doi: 10.1146/annurev-arplant-042809-112132
130. Bayer C, Vaupel P. Acute versus chronic hypoxia in tumors: Controversial data concerning time frames and biological consequences. *Strahlenther Onkol* (2012) 188(7):616–27. doi: 10.1007/s00066-012-0085-4
131. Rouschop KM, Ramaekers CH, Schaaf MB, Keulers TG, Savelkouls KG, Lambin P, et al. Autophagy is required during cycling hypoxia to lower production of reactive oxygen species. *Radiother Oncol* (2009) 92(3):411–6. doi: 10.1016/j.radonc.2009.06.029
132. Hsieh CH, Shyu WC, Chiang CY, Kuo JW, Shen WC, Liu RS. NADPH oxidase subunit 4-mediated reactive oxygen species contribute to cycling hypoxia-promoted tumor progression in glioblastoma multiforme. *PloS One* (2011) 6(9):e23945. doi: 10.1371/journal.pone.0023945
133. Kondo A, Safaei R, Mishima M, Niedner H, Lin X, Howell SB. Hypoxia-induced enrichment and mutagenesis of cells that have lost DNA mismatch repair. *Cancer Res* (2001) 61(20):7603–7.
134. Kato Y, Yashiro M, Fuyuhiro Y, Kashiwagi S, Matsuoka J, Hirakawa T, et al. Effects of acute and chronic hypoxia on the radiosensitivity of gastric and esophageal cancer cells. *Anticancer Res* (2011) 31(10):3369–75.

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 Vinaiphat, Low, Yeoh, Chng and Sze. 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 Oncology

ORIGINAL RESEARCH
published: 25 February 2021
doi: 10.3389/fonc.2021.540134

Quantitative Proteomic Profiling Identifies a Potential Novel Chaperone Marker in Resistant Breast Cancer

*Karen M. Gallegos[1], Jankiben R. Patel[1], Shawn D. Llopis[2], Rashidra R. Walker[1], A. Michael Davidson[1], Wensheng Zhang[3], Kun Zhang[3] and Syreeta L. Tilghman[1]**

[1] Division of Pharmaceutical Sciences, College of Pharmacy and Pharmaceutical Sciences, Florida A&M University, Tallahassee, FL, United States, [2] Division of Basic Pharmaceutical Sciences, College of Pharmacy, Xavier University of Louisiana, New Orleans, LA, United States, [3] Division of Mathematical and Physical Sciences, Department of Computer Science, College of Arts and Sciences, Xavier University of Louisiana, New Orleans, LA, United States

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Anca Maria Cimpean, Victor Babes University of Medicine and Pharmacy, Romania
Shiaw-Yih Lin, University of Texas MD Anderson Cancer Center, United States

****Correspondence:***
Syreeta L. Tilghman
syreeta.tilghman@famu.edu

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *03 March 2020*
Accepted: *06 January 2021*
Published: *25 February 2021*

Citation:
Gallegos KM, Patel JR, Llopis SD, Walker RR, Davidson AM, Zhang W, Zhang K and Tilghman SL (2021) Quantitative Proteomic Profiling Identifies a Potential Novel Chaperone Marker in Resistant Breast Cancer. Front. Oncol. 11:540134. doi: 10.3389/fonc.2021.540134

Development of aromatase inhibitor resistant breast cancer among postmenopausal women continues to be a major clinical obstacle. Previously, our group demonstrated that as breast cancer cells transition from hormone-dependent to hormone-independent, they are associated with increased growth factor signaling, enhanced cellular motility, and the epithelial to mesenchymal transition (EMT). Given the complexity of cancer stem cells (CSC) and their implications on endocrine resistance and EMT, we sought to understand their contribution towards the development of aromatase inhibitor resistant breast cancer. Cells cultured three dimensionally as mammospheres are enriched for CSCs and more accurately recapitulates tumors *in vivo*. Therefore, a global proteomic analysis was conducted using letrozole resistant breast cancer cells (LTLT-Ca) mammospheres and compared to their adherent counterparts. Results demonstrated over 1000 proteins with quantitative abundance ratios were identified. Among the quantified proteins, 359 were significantly altered ($p < 0.05$), where 173 were upregulated and 186 downregulated ($p < 0.05$, fold change >1.20). Notably, midasin, a chaperone protein required for maturation and nuclear export of the pre-60S ribosome was increased 35-fold. Protein expression analyses confirmed midasin is ubiquitously expressed in normal tissue but is overexpressed in lobular and ductal breast carcinoma tissue as well as ER+ and ER- breast cancer cell lines. Functional enrichment analyses indicated that 19 gene ontology terms and one KEGG pathway were over-represented by the down-regulated proteins and both were associated with protein synthesis. Increased midasin was strongly correlated with decreased relapse free survival in hormone independent breast cancer. For the first time, we characterized the global proteomic signature of CSC-enriched letrozole-resistant cells associated with protein synthesis, which may implicate a role for midasin in endocrine resistance.

Keywords: aromatase inhibitors, breast cancer, cancer stem cells, letrozole resistance, translation, midasin

INTRODUCTION

Aromatase inhibitors (AI), e.g., letrozole, are the first-line treatment for estrogen receptor positive (ER+) breast cancer in post-menopausal women. Despite widespread successful usage of letrozole, resistance to therapy, tumor relapse, and metastasis remain the principle causes of death for breast cancer patients (1, 2). While there are no cures for AI-resistant breast cancer, previous reports demonstrate AI resistance is associated with increased growth factor signaling (i.e., HER2) (3–5). In addition, we previously identified a global proteomic signature of letrozole resistance that was associated with enhanced cellular motility, estrogen independence (6) and the epithelial to mesenchymal transition (EMT) (7). Although the mortality rates for ER+ breast cancer in the US have declined due to successful endocrine therapy (AI and selective estrogen receptor modulators, SERMS), the development of resistance is still a critical lingering problem. In order to develop improved therapies for highly metastatic breast tumors, it is crucial to uncover the molecular underpinnings that drive resistance and proliferation. An exacerbating factor is the emergence of a small sub-population of breast cancer stem cells (CSC) in AI-resistant breast cancer (8). This is a rate-limiting factor as these pluripotent cells are *de novo* resistant to radiation and chemotherapy. While conventional therapeutic approaches decrease the tumor burden, those approaches are unable to target the CSC population, which usually drives tumor progression, invasion, and metastatic disease (9). To gain an insight of those molecular mechanisms, we identified a chaperone, midasin, as a potential marker for AI-resistant cancer in CSC-enriched mammospheres.

Malignant cancer cells including the CSC population proliferate continuously, thereby requiring enhanced translation including increased ribosome subunit synthesis; hence much of the cell energy is directed to the protein synthesis machinery (10). Ribosomes are the fundamental macromolecular machines at the core of translation. Formation of the 40S and 60S pre-ribosomal subunits are catalyzed by approximately 200 biogenesis factors that participate in the successive assembly and maturation steps, eventually leading to mature 40S and 60S ribosomal subunits (11–14). Among these are several energy-consuming enzymes including the midasin chaperone. (15–18). A homolog of midasin, Rea1, was initially identified in yeast (19, 20). It comprises an N-terminal domain, followed by six **A**TPase **a**ssociated with diverse cellular **a**ctivities (AAA) regions forming a ring, a linker domain, an acidic domain containing 35–40% aspartate and glutamate, and a carboxy-terminal M-domain that possesses Metal Ion Dependent Adhesion Site (MIDAS) sequence motifs and is homologous to the I-domain of integrins (15). Translation is a fine-tuned and tightly regulated process contributing to normal cell growth and development. It is well known that cancer cells have an increased demand for protein synthesis that is accomplished through aberrant changes in ribosome biogenesis (21). Likewise, inappropriate protein synthesis is associated with cancer. (22) Additionally, deregulation of protein synthesis could enable cells to acquire the classic hallmarks of cancer including, evading growth suppressors, sustaining proliferative signaling, activation of invasion and metastasis, and resisting cell death (23, 24). Although progress has been made towards demonstrating that changes in the quality and quantity of ribosomes can alter translation (25, 26), therapeutically targeting these ribosomal alterations has been unsuccessful.

Currently, effective targeted approaches to combat endocrine resistant breast cancer as well as other chemotherapy refractory cancers are lacking, in part due to an inability to inhibit CSCs and completely unravel the rate-limiting proteins and pathways that drive metastatic disease. To understand the contribution of CSCs on AI-resistance, we perform a quantitative proteomic analysis comparing letrozole-resistant breast cancer cells (LTLT-Ca) cultured two-dimensionally (2D) versus three-dimensionally (3D) as CSC-enriched mammospheres. We chose to use a 3D model system as 2D cultures do not represent the native events of breast cancer progression. Three-dimensional culture systems are advantageous because cells can be cultured in a spatially relevant manner that encourages cell and cell-matrix interactions that closely mimic the tumor environment and acquire morphological and cellular characteristics relevant to those of tumors *in vivo* (27). Since our previous study compared the proteomic profile of letrozole-sensitive vs letrozole-resistant cells cultured 2D, we were interested in capturing how the use of a more physiologically relevant model system would impact the proteome. Therefore, the study objective was to identify novel candidate proteins involved in endocrine resistance within the letrozole-resistant mammosphere cell population compared to cells cultured 2D.

MATERIALS AND METHODS

Cell Culture

Human LTLT-Ca cells are long-term letrozole treated MCF-7 cells and human AC-1 breast cancer cells are letrozole-sensitive MCF-7 cells. AC-1 cells were stably transfected with the human aromatase gene and the LTLT-Ca cells were derived from the AC-1 cells. Both were generously provided by the late Dr. Angela Brodie at University of Maryland. LTLT-Ca cells were cultured in 75-cm^2 flasks in phenol red-free IMEM (Invitrogen) supplemented with 10% charcoal-stripped fetal bovine serum (CS-FBS), 100 U/ml penicillin G sodium, 100 μg/ml streptomycin sulfate, and 0.25 μg/ml amphotericin B, 750 μg/ml geneticin (Invitrogen), and 1 μM letrozole (Sigma). AC-1 cells were cultured in 5% fetal bovine serum (FBS), 100 U/ml penicillin G sodium, 100 μg/ml streptomycin sulfate, and 0.25 μg/ml amphotericin B, and 750 μg/ml geneticin (Invitrogen). The culture flasks were maintained in a humidified atmosphere of 5% CO_2 at 37°C. The LTLT-Ca cells were isolated from tumors of aromatase transfected MCF-7 cells grown in ovariectomized severe combined immunodeficiency (SCID) mice following 56 weeks of treatment with letrozole as previously described (28). After long-term letrozole treatment, the tumors acquired the ability to proliferate in the presence of the drug. Mycoplasma testing has been performed on all cell lines. T47D letrozole-sensitive (T47Darom) and T47D letrozole-

resistant (T47DaromLR) cells were cultured as previously described by Gupta et al., (29). The MDA-MB-231 cell line (human breast cancer cells negative for ER, PR and Her2/neu) were acquired from the American Type Culture Collection (Manassas, VA, USA) and cultured as previously described (30). All cell lines were authenticated using Short Tandem Repeat (STR) analysis as described in 2012 in ANSI Standard (ASN-0002) by the ATCC Standards Development Organization and by Dr. L. Kerrigan's group (31). In brief, 17 STR loci plus the gender determining locus, Amelogenin, were amplified using the commercially available PowerPlex® 18D Kit (Promega). The cell line samples were processed using the ABI Prism® 3500 xL Genetic Analyzer. Data were analyzed using GeneMapper® ID-X v1.2 software (Applied Biosystems). The authentication results verified that AC-1 and LTLT-Ca cell lines, T47D variant cell lines, and the MDA-MB-231 cell line shared greater than 85% homology with the MCF-7 cell line, T47D cell line, and the MDA-MB-231 cell line, respectively. Cell lines with ® 80% match are considered to be related and derived from a common ancestry.

Mammosphere Formation Assay

LTLT-Ca cells were grown to 80–90% confluence and after media was removed, cells were rinsed twice with Hank's Balanced Salt Solution (HBSS) to remove residual culture media. Cells were gently scraped and resuspended in 10 ml of MammoCult™ media (Stemcell Technologies). Afterwards, cells were centrifuged at 500 g for 3 min at room temperature. The supernatant was discarded and the pellet was resuspended into a single cell suspension in 2 ml of MammoCult™ media. Cell concentration and viability was determined with Trypan Blue exclusion. Cells were enumerated and 100,000 cells were seeded in ultra-low adherent plates. The cultures were incubated in a 5% CO_2, humidified incubator at 37° C for 7 days and spheres greater than 60 μm were counted and recorded. Mammosphere formation was identified by light microscopy and then harvested as indicated in western blot analysis.

Western Blotting Analysis

2D and 3D cultured cells were homogenized in cold RIPA buffer supplemented with 2x protease and phosphatase inhibitors (ThermoFisher Scientific). The supernatant was incubated with Laemmli protein sample buffer (Bio-Rad) at 70°C for 10 min. About 75 μg of denatured protein was separated on 7.5% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad) and transferred to PVDF membranes. All blots were blocked for 1 h with 5% Bovine Serum Albumin (BSA) in Phosphate-Buffered Saline, 0.1% Tween (PBS-T) buffer. Following incubation with anti-midasin antibody (Sigma Aldrich) and secondary antibodies, the blots were visualized using ChemiDoc XRS imaging system (BioRad) and detected with the Clarity Max Western ECL Substrate (BioRad). The exposure time was automatically detected by the imaging system. The protein bands were analyzed using Image Lab software (BioRad). Arbitrary densitometry units were quantified and expressed as mean ± standard deviation. Midasin protein expression was normalized to housekeeping protein bands (GAPDH). All experiments were performed with $n \geq 3$ and a total of 3 biological replicates were performed.

Immunohistochemistry

The immunostaining protocol for ICC can be found on the open access repository for science methods at: https://www.protocols.io/view/hpa-cell-atlas-standard-immunostaining-protocol-x2dfqa6.

To provide an overview of protein expression patterns, all images of tissues stained by immunohistochemistry were manually annotated by a specialist followed by verification by a second specialist. Annotation of each different normal and cancer tissue is performed using fixed guidelines for classification of immunohistochemical results. Each tissue is examined for representability, and subsequently immunoreactivity in the different cell types present in normal or cancer tissues was annotated. Basic annotation parameters include an evaluation of i) staining intensity (negative, weak, moderate or strong), ii) fraction of stained cells (<25%, 25%–75% or >75%) and iii) subcellular localization (nuclear and/or cytoplasmic/membranous). The scanned images were viewed in 20x magnification in the freely available database (www.proteinatlas.org). If there was any discordance between the observers, a shared review of the images was performed to obtain a common interpretation.

Cell Lysis for Proteomic Analysis

LTLT-Ca cells were cultured to 80% confluence in the medium as described above, and washed with cold HBSS three times, then collected with cell scraper. NP40 cell lysis buffer (Invitrogen) containing additional 1 mM of phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (Sigma) was used to extract total cellular proteins. The concentration of proteins was measured with BCA assay (Pierce Biotechnology, Rockford, IL). The cell lysates were stored at -80 °C before further processing. There was a total of 3 biological replicates for each sample.

Trypsin Digestion for Proteomic Analysis

Protein samples were digested with sequencing grade modified trypsin (Promega Corp) according to manufacturer's instructions. Briefly, 45 μl of 200 mM triethyl ammonium bicarbonate (TEAB) was added to aliquots of 100 μg of protein sample and the final volume was adjusted to 100 μl with ultrapure water. A total of 5 μl of 200 mM Tris (2-carboxyethyl) phosphine (TCEP) was added and the resulting mixture was incubated for 1 h. Following the addition of 5 μl of 375 mM iodoacetamide, the mixture was incubated for 30 min in the dark. After incubation, 1 ml of pre-chilled acetone was added and the precipitation was allowed to proceed overnight. The acetone-precipitated protein pellets were suspended with 100 μl of 200 mM TEAB and 2.5 μg of trypsin was added to digest the sample overnight at 37° C.

Tandem Mass Tags (TMT) Labeling for Proteomic Analysis

Tandem mass tags TMT6 (Thermo Scientific) with different molecular weights (126 ~ 131 Da) were applied as isobaric tags for relative and absolute quantification during mass spectrometry

analysis. According to the manufacturer's protocols, the digested samples were individually labeled with TMT6 reagents for 1h as follows: three 100-µg aliquots of digested peptides from LTLT-Ca adherent cells were each labeled with a different isobaric tag (TMT126, 127, and 128, respectively). Likewise, 100-µg aliquots of peptides from LTLT-Ca mammospheres were labeled with TMT129, 130, and 131 mass tags, respectively. The labeling reaction was quenched with 5% hydroxylamine. Finally, the six labeled peptide aliquots were combined for subsequent fractionation.

Fractionation of Labeled Peptide Mixture Using a Strong Cation Exchange Column

The combined TMT labeled peptide mixture was fractionated with a strong cation exchange column (SCX) (Thermo Scientific) on a Shimadzu 2010 HPLC equipped with a UV detector (Shimadzu). Mobile phase consisted of buffer A (5 mM KH_2PO_4, 25% acetonitrile, pH 2.8) and buffer B (buffer A plus 350 mM KCl). The column was equilibrated with Buffer A for 30 min before sample injection. The mobile phase gradient was set as follows at a flow rate of 1.0 ml/min: (a) 0 to 10 min: 0% buffer B; (b) 10 to 40 min: 0% to 25% Buffer B, (c) 40 to 45 min: 25% to 100% Buffer B; (d) 45 to 50 min: 100% buffer B; (e) 50 to 60 min: 100% to 0% buffer B; (f) 60 min to 90 min: 0% buffer B. A total of 60 fractions were initially collected, combined and lyophilized into 15 final fractions based on SCX chromatographic peaks.

Desalination of Fractionated Samples for Proteomic Analysis

A C_{18} solid-phase extraction (SPE) column (Hyper-Sep SPE Columns, Thermo-Fisher Scientific) was used to desalt all collected fractions. The combined 15 fractions were each adjusted to a final volume of 1 ml containing 0.25% (v/v) trifluoroacetic acid (TFA, Sigma). The C_{18} SPE columns were conditioned before use by filling them with 1 ml acetonitrile and allowing the solvent to pass through the column slowly. The columns were then rinsed three times with 1 ml 0.25% TFA solution. The fractions were loaded on top of the SPE cartridge and allowed to elute slowly. Columns were washed four times with 1 ml 0.25% TFA aliquots before the peptides were eluted with 3x 400 µl of 80% acetonitrile/0.1% formic acid.

Liquid Chromatography Linked to Tandem Mass Spectrometry (LC–MS/MS) Analysis on LTQ-Orbitrap

Peptides were analyzed on an LTQ-Orbitrap XL instrument (Thermo-Fisher Scientific) coupled to an Ultimate 3000 Dionex nanoflow LC system (Dionex). High mass resolution was used for peptide identification and high energy collision dissociation (HCD) was employed for reporter ion quantification. The RP-LC system consisted of a peptide Cap-Trap cartridge (0.5 x 2 mm) (Michrom BioResources) and a prepacked BioBasic C_{18} PicoFrit analytical column of 75 µm i.d. × 15 cm length (New Objective) fitted with a FortisTip emitter tip. Samples were loaded onto the trap cartridge and washed with mobile phase A (98% H_2O, 2% acetonitrile and 0.1% formic acid) for concentration and desalting. Subsequently, peptides were eluted over 180 min from the analytical column *via* the trap cartridge using a linear gradient of 6%–100% mobile phase B (20% H_2O, 80% acetonitrile and 0.1% formic acid) at a flow-rate of 0.3 µl/min using the following gradient: 6% B for 5 min; 6%–60% B for 125 min; 60%–100% B for 5 min; hold at 100% B for 5 min; 100%–6% B in 2 min; hold at 6% B for 38 min.

The LTQ-Orbitrap tandem mass spectrometer was operated in a data-dependent mode. Briefly, each full MS scan (60,000 resolving power) was followed by six MS/MS scans where the three most abundant molecular ions were dynamically selected and fragmented by collision-induced dissociation (CID) using a normalized collision energy of 35%; the same three molecular ions were also scanned three times by HCD-MS^2 with collision energy of 45%. MS scans were acquired in profile mode and MS/MS scans in centroid mode. LTQ-Orbitrap settings were as follows: spray voltage 2.0 kV, 1 microscan for MS1 scans at 60,000 resolution (fwhm at m/z 400), microscans for MS^2 at 7500 resolution (fwhm at m/z 400); full MS mass range, m/z 400 −1400; MS/MS mass range, m/z 100–2000. The "FT master scan preview mode", "Charge state screening", "Monoisotopic precursor selection", and "Charge state rejection" were enabled so that only the 2+, 3+ and 4+ ions were selected and fragmented by CID and HCD.

Gene Ontology and KEGG Analysis of Proteomics Data

Functional enrichment analysis were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (32) which is a resource consisting of an integrated biological knowledgebase and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists. Using this tool, the proteomic profile of the LTLT-Ca mammospheres versus the LTLT-Ca adherent cells were analyzed and used to identify the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) terms over-represented by the significantly expressed proteins (genes).

Molecular Pathway and Network Analysis in Ingenuity Pathway Analysis Software

To systematically evaluate the differences between the LTLT-Ca 3D and LTLT-Ca 2D cells, the proteomic signatures of the two groups were subjected to pathway analysis using IPA software (Ingenuity Systems). Accession numbers and fold changes of proteins from Student's *t*-tests were imported into IPA to identify the biological relationships among the proteins. Canonical pathways and interaction networks were generated based on the knowledge sorted in the Ingenuity Pathway Knowledge base.

Kaplan Meier (KM) Survival Analysis

The application of KM plot has been described in detail previously (33, 34). Briefly, KM plots were obtained using the KM Plotter web-based (kmplot.com/analysis) curator, which surveys public microarray repositories for relapse free and overall survival among patients with breast, lung, ovarian or gastric cancers. The KM Plotter recognizes 54,675 individual

Affymetrix probe sets, and surveys expression data from 4,142 breast cancer patients (as of 2014). Survival and gene expression data were derived from the GEO (Gene Expression Omnibus), TCGA (The Cancer Genome Atlas), and EGA (European Genome-phenome Atlas) databases. Patients were stratified as ER+ or ER-. In order to ascertain midasin expression, Affymetrix probe 212693 was selected. Populations were split by median ERα expression and plots generated accordingly. Relapse Free Survival (RFS) in the total population (2557 patients) was determined.

Statistical Analysis

Results are expressed as the mean unit ± standard error of the mean (SEM) (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$) using the Graph Pad Prism V.6 software program.

RESULTS

Quantitative Proteomic Analysis Reveals Extensive Changes in Protein Expression in Letrozole-Resistant Mammospheres (3D) Compared to Letrozole-Resistant Adherent Cells (2D)

In order to characterize the proteome of LTLT-Ca cells cultured as CSC-enriched mammospheres (3D) and compare it to their LTLT-Ca adherent cell (2D) counterparts, we performed a gel-free proteomic analysis. This approach combined TMT labeling, two-dimensional HPLC, and high-resolution MS. Over 1,000 proteins with quantitative abundance ratios were identified. Among the quantified proteins, 359 were significantly altered ($p < 0.05$), where 173 were up regulated (**Table 1**) and 186 down regulated (**Table 2**) ($p < 0.05$, fold change >1.20). The most significantly changed protein was a 35-fold increase in midasin, a nuclear chaperone protein required for maturation and nuclear export of the pre-60S ribosomal subunit (15, 17, 35). To validate this finding, western blotting analyses were conducted, and results confirmed that the LTLT-Ca mammospheres (3D) expressed significantly higher levels of midasin compared to the LTLT-Ca adherent cells (2D) (**Figures 1A, B**). A similar result was also observed when letrozole-sensitive AC-1 mammospheres (3D) were compared to letrozole-sensitive AC-1 adherent cells 2D (**Figure 1**). To examine whether increased midasin expression was exclusive to the LTLT-Ca mammospheres, we investigated midasin expression across various cell lines including letrozole-sensitive and letrozole-resistant cell lines. When the 3D mammospheres were compared to their 2D culture counterparts, midasin expression was significantly increased in the letrozole resistant T47D breast cancer cells (T47DaromLR) (**Figure S1A, B**) as well as the MDA-MB-231 triple negative breast cancer cells (**Figure S1C, D**). The MDA-MB-231 mammospheres exhibited an increase in midasin expression greater than 1000% compared to their 2D counterparts which was the most dramatic change in midasin expression. Interestingly, the T47Darom mammospheres expressed significantly lower midasin expression compared to the T47Darom 2D cultures.

Additionally, we observed increased expression of S100-P, Profilin-1, Elongation factor 1α1, and serotransferrin. In our previous report, letrozole resistant cells (LTLT-Ca) were compared to letrozole-sensitive cells (AC-1) and a global proteomic signature was identified where several proteins were increased in the letrozole-resistant LTLT-Ca adherent 2D cells (ie. Protein S100-P and Pofilin-1) versus AC-1 adherent 2D cells (6). Here, those proteins are further increased in the LTLT-Ca 3D mammospheres versus 2D adherent cells.

Letrozole Resistant Mammospheres Are Associated With Increased Components of the Translational Machinery

An additional analysis of the proteomes was performed using the Benjamini-Hochberg (BH) method. The selection was based on the following criteria: (1) the adjusted p value < 0.01 (corrected using the BH method) (36) and (2) the average expression ratio greater than 1.5 for upregulated proteins or less than 0.67 for down-regulated proteins. Using the more stringent BH method, we identified 55 upregulated proteins (**Table S1**) and 65 downregulated proteins (**Table S2**) in the 3D LTLT-Ca mammospheres compared to the 2D adherent cells. Using this data, functional enrichment analyses were conducted utilizing the DAVID tool (http://david.abcc.ncifcrf.gov/) (32). Based on the result, 19 GO terms (**Figure 2** and **Table S3**) and one KEGG pathway (hsa03010:Ribosome) were over-represented (BH adjusted p value < 0.01) by the top 65 downregulated (BH adjusted p-value < 0.01) proteins. The KEGG term was enriched 29-fold (**Table S4**). Several of these GO terms including "translational elongation", "ribosomal subunit", "cytosolic ribosome" and "translation" was enriched by over 15-fold and of the 19 GO terms identified, 15 were related to protein synthesis and ribosomes. Interestingly, the upregulated protein set was less informative. Only one cellular component GO term "cytosol" and one biological process GO term "response to organic cyclic substance" were over-represented by the proteins. Our findings of increased midasin expression coupled with the GO and KEGG analyses, supported the previously identified role of midasin. Further, many of the significantly altered proteins shown in **Table 2** were ribosomal proteins associated with the 40S and 60S ribosomes (highlighted with red stars in **Figure 3**). While several of the 40S and 60S ribosomal proteins were significantly altered, more of the 40S ribosomal proteins were decreased compared to the 60S ribosomal proteins (**Table S2**). The downregulated proteins associated with the 40S ribosomal subunit included, RPS9, RPS10, and RPS14, among others, while those associated with the 60S subunit included RPL11, RPL10a, and RPL7a. Taken together, these findings suggest increased midasin expression in CSC-enriched letrozole-resistant mammospheres is highly associated with alterations in the translational machinery impacting the composition of ribosomal proteins, which may contribute to endocrine resistance.

To develop a network view of predicted protein associations, a molecular pathway and network analysis using IPA software was used with the differentially regulated proteins identified by MS

TABLE 1 | Proteomic analyses of selected up-regulated proteins in LTLT-Ca letrozole-resistant mammospheres where ratios of the fold change (3D/2D) and *p* values are shown.

Accession	Description	# AAs	MW [kDa]	Fold change of LTLT-Ca (3D/2D)	*t*-test (*p* values)
IPI00167941.1	Midasin	5596	632.4	35.049	5.60E-05
IPI00022463.2	Serotransferrin	698	77.0	10.988	3.22E-05
IPI00871227.1	Isoform 1 of Hemicentin-1	5635	613.0	8.160	3.95E-04
IPI00942032.1	Anterior gradient protein 2 homolog	175	20.0	7.932	6.77E-08
IPI00073772.6	Fructose-1,6-bisphosphatase 1	338	36.8	3.974	8.18E-05
IPI00976367.1	Uncharacterized protein	76	9.0	3.408	3.07E-05
IPI00028376.1	Mitochondrial import inner membrane translocase subunit Tim8 A	97	11.0	2.971	3.14E-04
IPI00465352.2	Calcyphosin	189	21.0	2.626	2.47E-04
IPI00017526.1	Protein S100-P	95	10.4	2.717	1.04E-03
IPI00301058.5	Vasodilator-stimulated phosphoprotein	380	39.8	2.382	5.95E-05
IPI00646689.1	Thioredoxin domain-containing protein 17	123	13.9	2.482	5.28E-05
IPI00292020.3	Spermidine synthase	302	33.8	2.306	3.35E-06
IPI00785061.1	Isoform 2 of Membrane metallo-endopeptidase-like 1	611	69.9	2.173	3.89E-04
IPI00106687.4	Latexin	222	25.7	2.188	8.40E-05
IPI00304409.3	Calcium-regulated heat stable protein 1	147	15.9	2.344	2.06E-03
IPI00010182.4	Isoform 1 of Acyl-CoA-binding protein	87	10.0	2.335	3.99E-03
IPI00745466.1	Isoform ARPP-16 of cAMP-regulated phosphoprotein 19	96	10.6	2.141	4.50E-03
IPI00012197.1	dCTP pyrophosphatase 1	170	18.7	1.870	2.45E-03
IPI00964927.1	UDP-glucose 6-dehydrogenase isoform 3	397	44.4	1.921	1.58E-04
IPI00003935.6	Histone H2B type 2-E	126	13.9	1.865	1.26E-04
IPI00418262.5	Fructose-bisphosphate aldolase	451	48.4	2.181	7.56E-04
IPI00916111.3	Malate dehydrogenase, cytoplasmic	334	36.4	1.862	4.57E-05
IPI00021828.1	Cystatin-B	98	11.1	1.978	1.69E-03
IPI00883944.1	Gamma-synuclein	79	8.3	1.690	1.98E-03
IPI00953696.1	glutathione reductase, mitochondrial isoform 4 precursor	440	47.2	1.809	8.69E-05
IPI00301434.4	BolA-like protein 2	152	16.9	1.604	9.27E-03
IPI00843765.1	Isoform 3 of Spectrin alpha chain, brain	2452	282.1	1.798	1.14E-03
IPI00853358.1	MUC1 isoform T9	110	12.1	1.616	7.89E-04
IPI00062037.1	Dynein light chain 2, cytoplasmic	89	10.3	1.745	1.42E-06
IPI00218829.9	Eukaryotic peptide chain release factor GTP-binding subunit ERF3A	499	55.7	1.647	1.27E-04
IPI00827491.1	Isoform 2 of Ubiquitin-like modifier-activating enzyme 6	578	65.7	1.711	5.60E-05
IPI01012504.1	6-phosphogluconate dehydrogenase, decarboxylating	470	51.8	1.543	3.61E-03
IPI00426077.4	Isoform 2 of Calcineurin-like phosphoesterase domain-containing protein 1	172	19.0	1.737	4.85E-05
IPI00465439.5	Fructose-bisphosphate aldolase A	364	39.4	1.683	1.55E-04
IPI00414315.10	Isoform 1 of Epidermal growth factor receptor kinase substrate 8-like protein 2	715	80.6	1.561	5.92E-04
IPI00027933.3	Proteasome subunit beta type-10	273	28.9	1.620	2.83E-03
IPI00219446.5	Phosphatidylethanolamine-binding protein 1	187	21.0	1.677	1.54E-04
IPI00419585.9	Peptidyl-prolyl cis-trans isomerase A	165	18.0	1.585	2.12E-06
IPI00007052.6	Mitochondrial fission 1 protein	152	16.9	1.536	4.54E-05
IPI00979393.2	Isoform 3 of Calpastatin	590	63.6	1.607	4.78E-03
IPI00027223.2	Isocitrate dehydrogenase [NADP] cytoplasmic	414	46.6	1.515	6.23E-05
IPI00396485.3	Elongation factor 1-alpha 1	462	50.1	1.460	1.93E-03
IPI00411704.9	Isoform 1 of Eukaryotic translation initiation factor 5A-1	154	16.8	1.565	1.79E-03
IPI00219953.5	UMP-CMP kinase isoform a	228	25.8	1.512	1.68E-03
IPI00794609.1	Cysteine synthase	177	19.1	1.457	1.32E-04
IPI00015018.1	Inorganic pyrophosphatase	289	32.6	1.514	1.24E-04
IPI00398922.4	Protein phosphatase 1 regulatory subunit 14B	199	21.0	1.564	9.47E-03
IPI00220637.5	Seryl-tRNA synthetase, cytoplasmic	514	58.7	1.412	4.43E-03
IPI00221362.3	Isoform 3 of Apoptosis-associated speck-like protein containing a CARD	135	15.0	1.337	7.02E-03
IPI00815656.1	Parathymosin (Fragment)	63	7.4	1.700	2.33E-02
IPI00903145.1	Radixin	583	68.5	1.410	3.09E-04
IPI00295386.7	Carbonyl reductase [NADPH] 1	277	30.4	1.367	5.76E-03
IPI00025285.3	V-type proton ATPase subunit G 1	118	13.7	1.250	1.03E-01
IPI00028091.3	Actin-related protein 3	418	47.3	1.397	1.51E-03
IPI00479186.7	Isoform M2 of Pyruvate kinase isozymes M1/M2	531	57.9	1.428	5.13E-06
IPI00028004.2	Proteasome subunit beta type-3	205	22.9	1.289	1.32E-02
IPI00797270.4	Isoform 1 of Triosephosphate isomerase	249	26.7	1.402	1.10E-05
IPI00375513.3	Isoform Soluble of Catechol O-methyltransferase	221	24.4	1.513	1.99E-03
IPI00294398.2	Isoform 1 of Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	314	34.3	1.270	9.69E-03
IPI00465436.4	Catalase	527	59.7	1.353	1.84E-04
IPI00219018.7	Glyceraldehyde-3-phosphate dehydrogenase	335	36.0	1.357	1.54E-04

(Continued)

TABLE 1 | Continued

Accession	Description	# AAs	MW [kDa]	Fold change of LTLT-Ca (3D/2D)	*t*-test (*p* values)
IPI00216691.5	Profilin-1	140	15.0	1.319	8.52E-04
IPI00298547.3	Protein DJ-1	189	19.9	1.410	4.13E-03
IPI01011912.1	Phosphoglycerate kinase	389	41.4	1.330	8.14E-05
IPI00026781.3	Fatty acid synthase	2511	273.3	1.340	1.44E-03
IPI00217966.9	Isoform 1 of L-lactate dehydrogenase A chain	332	36.7	1.276	4.72E-04
IPI01011075.1	Superoxide dismutase	162	18.3	1.386	3.51E-03
IPI00479145.3	Keratin, type I cytoskeletal 19	400	44.1	1.364	3.25E-04
IPI00016610.2	Poly(rC)-binding protein 1	356	37.5	1.337	1.58E-02
IPI00011229.1	Cathepsin D	412	44.5	1.254	1.33E-03
IPI00218070.1	Isoform GRB3-3 of Growth factor receptor-bound protein 2	176	20.5	1.505	6.20E-03
IPI00645307.2	Isoform 1 of Isopentenyl-diphosphate Delta-isomerase 1	227	26.3	1.303	1.55E-03
IPI00218482.2	Isoform Short of ES1 protein homolog, mitochondrial	237	24.7	1.297	1.10E-03
IPI00011416.2	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	328	35.8	1.389	7.22E-03
IPI00014808.1	Platelet-activating factor acetylhydrolase IB subunit gamma	231	25.7	1.240	5.28E-03
IPI00028109.1	Protein dpy-30 homolog	99	11.2	1.257	4.62E-03
IPI00908881.3	Glucose-6-phosphate isomerase	531	60.0	1.241	8.26E-03
IPI00010896.3	Chloride intracellular channel protein 1	241	26.9	1.257	1.66E-04
IPI00147874.1	Sialic acid synthase	359	40.3	1.330	5.10E-04
IPI00218493.7	Hypoxanthine-guanine phosphoribosyltransferase	218	24.6	1.193	6.98E-02
IPI00981623.1	Uncharacterized protein	98	10.6	1.215	1.97E-02
IPI00377066.1	ras suppressor protein 1 isoform 2	224	25.5	1.209	5.05E-02
IPI00472003.1	Isoform 4 of BH3-interacting domain death agonist	99	11.3	1.268	3.76E-04
IPI00216770.1	Isoform 2 of 26S protease regulatory subunit 6B	387	43.5	1.131	2.17E-01
IPI00550069.3	Ribonuclease inhibitor	461	49.9	1.261	8.96E-07

TABLE 2 | Proteomic analyses of selected down-regulated proteins in LTLT-Ca letrozole-resistant mammospheres where ratios (3D/2D) of the fold changes and *p* values are shown.

Accession	Description	# AAs	MW [kDa]	Fold change (3D/2D)	*t*-test (*p* values)
IPI00013485.3	40S ribosomal protein S2	293	31.3	-1.704	7.75E-05
IPI00012750.3	40S ribosomal protein S25	125	13.7	-1.780	4.94E-06
IPI00221091.9	40S ribosomal protein S15a	130	14.8	-1.784	2.22E-04
IPI00216514.6	Isoform OA3-293 of Leukocyte surface antigen CD47	292	31.7	-1.784	4.89E-05
IPI00647168.1	Ribosomal protein L11	131	14.9	-1.703	1.29E-04
IPI00215780.5	40S ribosomal protein S19	145	16.1	-1.782	2.92E-04
IPI00026271.5	40S ribosomal protein S14	151	16.3	-1.746	6.69E-05
IPI00908521.1	Isoform 2 of Brain acid soluble protein 1	173	17.7	-1.673	2.78E-03
IPI00644936.1	Isoform 3 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms short	379	44.2	-2.055	2.63E-05
IPI00926712.1	Integrin beta	681	74.9	-2.021	3.96E-08
IPI00011253.3	40S ribosomal protein S3	243	26.7	-1.849	1.43E-04
IPI00019472.4	Neutral amino acid transporter B(0)	541	56.6	-2.039	8.05E-06
IPI00979226.1	60S ribosomal protein L17 isoform b	146	17.1	-2.004	8.59E-05
IPI00337335.8	Isoform 1 of Myosin-14	1995	227.7	-1.967	7.32E-05
IPI00922108.2	integrin alpha-V isoform 2	1002	111.1	-2.111	8.17E-06
IPI00645589.1	Isoform 3 of Choline transporter-like protein 2	704	79.8	-2.089	5.47E-05
IPI00985384.1	ATP-dependent RNA helicase DDX3X isoform 3	646	71.3	-2.194	6.26E-05
IPI00008438.1	40S ribosomal protein S10	165	18.9	-2.337	5.48E-07
IPI00396321.1	Leucine-rich repeat-containing protein 59	307	34.9	-2.144	5.22E-05
IPI00015842.1	Reticulocalbin-1	331	38.9	-2.136	2.76E-04
IPI00221088.5	40S ribosomal protein S9	194	22.6	-2.467	7.90E-04
IPI00455315.4	Isoform 1 of Annexin A2	339	38.6	-2.496	5.64E-05
IPI00009865.4	Keratin, type I cytoskeletal 10	584	58.8	-2.176	7.64E-02
IPI00297910.2	Tumor-associated calcium signal transducer 2	323	35.7	-3.207	9.52E-06

(**Figure 4**). The resulting network demonstrated that midasin has a direct interaction with the ubiquitin thioesterase (OTUB1) with a 1.13-fold increase. This finding was interesting as OTUB1 inhibits the degradation of FOXM1 transcription factor, typically upregulated and overexpressed in aggressive therapy resistant breast cancer (37). Additionally, midasin directly interacted with glutamyl-prolyl-tRNA synthetase (EPRS), with a -1.23-fold decrease. EPRS is responsible for charging tRNAs with their cognate amino acids and is reported in the regulation of breast tumorigenesis (38). On the other hand, midasin indirectly

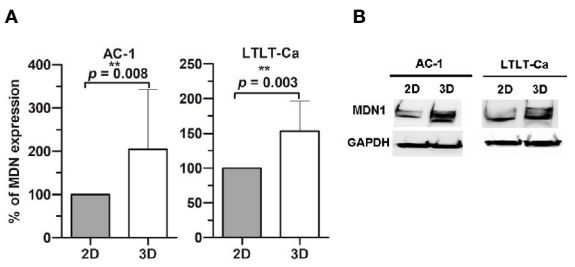


FIGURE 1 | Western Blot analysis of midasin expression. Midasin expression in letrozole-sensitive (AC-1) and letrozole-resistant (LTLT-Ca) breast cancer cells cultured adherently (2D) or as mammospheres (3D). All cells were evaluated by immunoblot to examine the expression of midasin and GAPDH (loading control). **(A)** Graphs depict normalized percentages of protein expression intensities relative to 2D cell counterparts. **(B)** Representative immunoblot depicts the protein expression of midasin and GAPDH.

interacted with calcium-regulated heat stable protein 1 (CARHSP1) and phenylalanyl-tRNA synthetase subunit alpha (FARSA) through direct interactions with ubiquitin C (UBC). CARHSP1 binds and regulates the stability of target mRNAs and increased by 2.34-fold, while FARSA is involved in charging tRNAs with their cognant amino acids and decreased by 2.14-fold. These direct and indirect protein interactions with midasin shed additional light on the inherent complexity of the proteome of CSC-enriched letrozole-resistant breast cancer.

Increased Midasin Expression in Breast Cancer Tissue

Previously, Uhlén et al., mapped the human tissue proteome using quantitative transcriptomics at the tissue and organ level, combined with tissue microarray-based immunohistochemistry (39). The data was integrated into an interactive web-based database allowing exploration of individual proteins and global expression patterns in all major tissues and organs of the human body. The database serves as a repository for immunohistochemistry tissue samples from breast cancer patients. This database was mined to further explore the expression pattern of midasin in normal breast tissue (**Table 3**) and breast cancer tissue samples (**Table 4**). In normal breast tissue, the adipocytes, glandular, and myoepithelial cells were examined for midasin expression (**Figures 5A, B**) and in most instances the tissues stained weakly for midasin. Likewise, when lobular carcinoma tissue samples were examined, the majority of the samples exhibited weak staining intensity while one sample expressed moderate staining intensity (**Figures 5C–F**). However, the majority of ductal carcinoma tissue samples expressed moderate staining intensity (**Figures 5G–J**). Regardless of the tissue type examined (ie., normal or cancerous), midasin was predominately localized to the cytoplasm and membrane. As such, the immunohistochemistry analyses confirmed that midasin is ubiquitously expressed in both normal breast tissue and cancer tissue, but ductal carcinoma tumor tissue had a higher midasin staining intensity.

High Midasin Expression Is Positively Correlated With Decreased RFS in ER- Breast Cancer Patients

As immunohistochemistry results demonstrated that some breast cancer tissue expressed higher levels of midasin compared to normal breast tissue samples, it was important to examine the clinical relevance of midasin levels on patient survival and determine within this context whether hormone receptor status

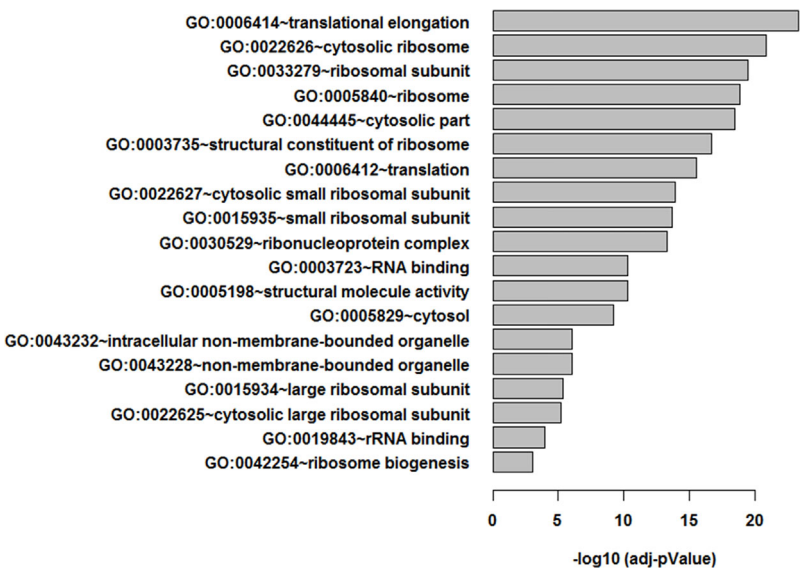


FIGURE 2 | Functional enrichment analysis of LTLT-Ca proteome. Functional enrichment analysis were performed using the DAVID tool (32), resulting in 19 gene ontology (GO) terms.

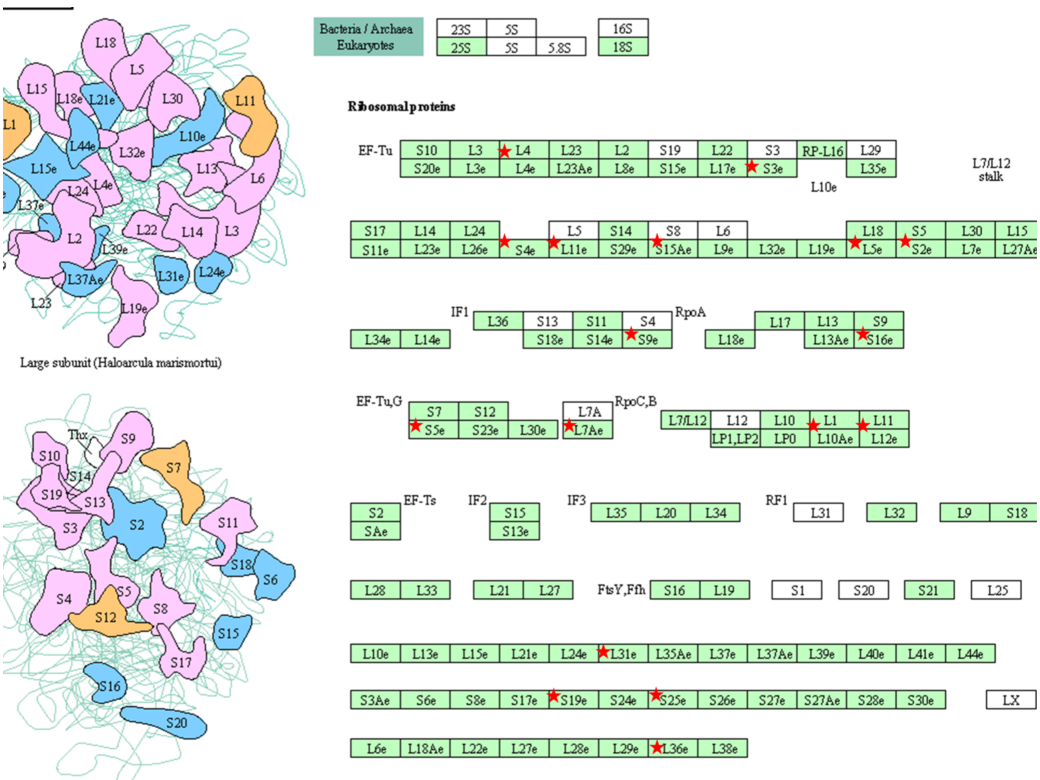


FIGURE 3 | KEGG pathway (hsa03010:Ribosome) Graphs. KEGG pathway were over-represented (BH adjusted *p* value < 0.01) by the top 65 down-regulated (BH adjusted p-value < 0.01) proteins. The pathway hsa03010:Ribosome was graphically represented and the ribosomal proteins were marked with red stars. The graphics were retrieved from the output of the functional enrichment analysis using the DAVID tool. The bar plot was generated by our lab-owned R codes.

was relevant to patient outcomes. To predict whether midasin expression was associated with differences in RFS, KM Plotter was used to interrogate publicly available microarray repositories for ER+ and ER- breast cancer patients. We chose to stratify patients based on hormone receptor status because our LTLT-Ca model of letrozole resistance is hormone independent (40) and reflects patients within this population as they progress to metastatic disease. Based on these parameters, high midasin expression was associated with a significant decrease in RFS among ER- patients ($p = 0.0214$) (**Figure 6A**) whereas increased midasin expression did not result in a statistically significant difference in RFS among ER+ patients (**Figure 6B**). Altogether, we demonstrate for the first time, a global proteomic signature for cancer stem cell enriched letrozole resistant mammospheres that is associated with high midasin expression, increased components of the translational machinery, and finally, that increased midasin expression is associated with decreased RFS in ER- breast cancer patients.

DISCUSSION

Overcoming aromatase inhibitor resistance continues to be a challenge clinically, further compounded by the radiation and chemotherapy resistant nature of breast CSCs. It has been shown that culturing cells as mammospheres allows for the propagation of breast cancer stem cells (41). This is critical as 3D cultures are more physiologically relevant than 2D cultures and mammospheres and adherent cells possessed markedly different proteomes (42). To this end, whole MCF-7 mammospheres, as well as subpopulations within spheres, have been shown to be more tumorigenic than 2D monolayer parental cultures (43–45). This suggests an enriched population of breast CSCs within spheres may be partly responsible for conferring tumorigenesis. To gain a deeper understanding of the mechanism that drives AI-resistant breast cancer, previous studies from our lab compared the proteome of letrozole-sensitive AC-1 breast cancer cells to letrozole-resistant LTLT-Ca breast cancer cells, where both cell lines were cultured adherently (2D). We identified a global proteomic signature associated with hormone independence, enhanced cell motility, and EMT among the 1,743 differentially regulated proteins (6). However, one of the limitations of this previous study was evaluating cells cultured as monolayers (2D) may have prevented the identification of proteins expressed when cells are cultured 3D as CSC-enriched mammospheres. Therefore, in this current study, we evaluated the proteome of the letrozole-resistant CSC-enriched mammosphere population and compared it to letrozole-resistant adherent cells to identify additional candidate proteins involved in endocrine resistance.

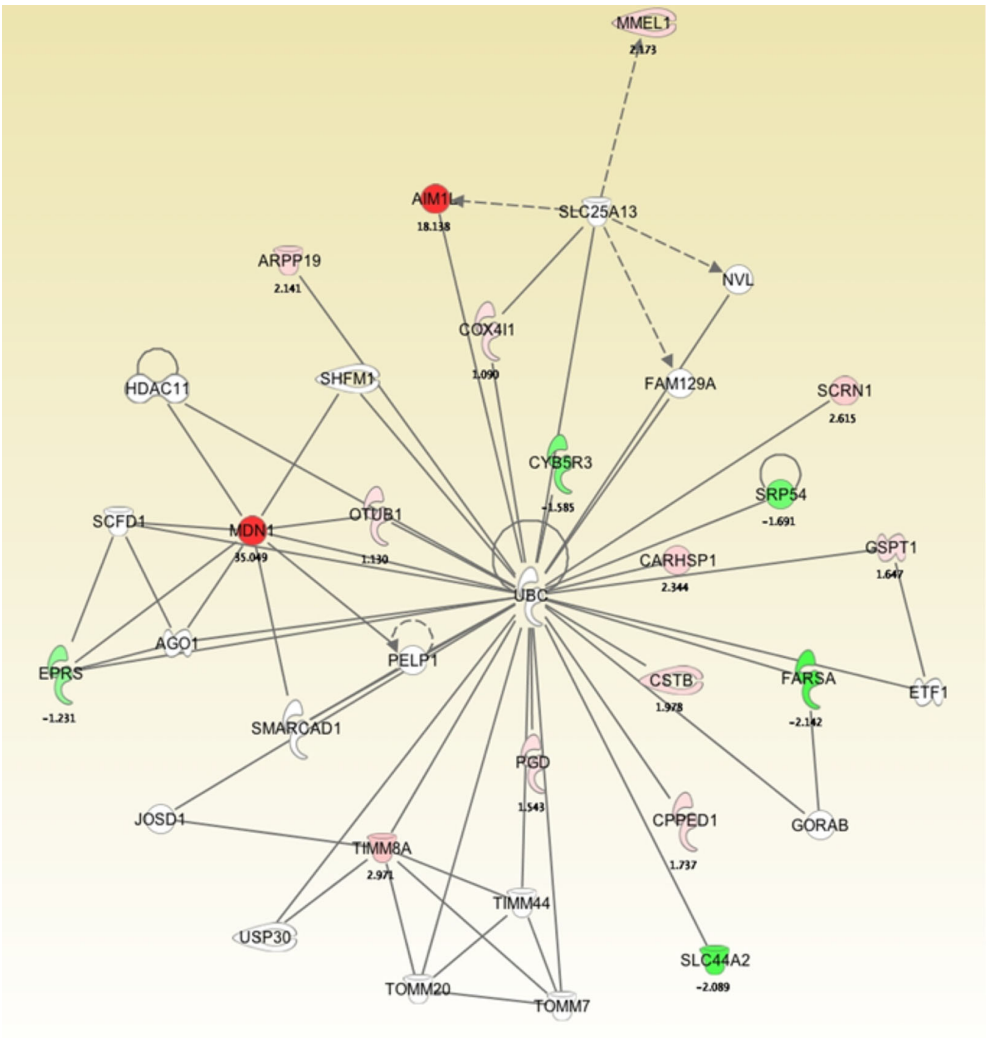


FIGURE 4 | Molecular pathway and network analysis of proteomics analysis of differentially expressed proteins in LTLT-Ca mammospheres and LTLT-Ca adherent cells. The proteome data was uploaded into IPA software that determined a network view of predicted associations for midasin. The network nodes represent proteins, and solid lines indicated a direct relationship while dashed lines represent an indirect relationship between proteins.

Here, we identified a novel proteomic signature associated with translation that stemmed from the observation that midasin was upregulated by 35-fold in LTLT-Ca mammospheres (**Table 1**, **Figure 1**). This was a unique feature of this signature as midasin was not detected when the 2D LTLT-Ca cells were compared to 2D AC-1 cells (6). The present finding of increased midasin expression in cancer stem cell enriched mammospheres is highly relevant, as midasin plays a key role in driving nuclear export of pre-ribosomal particles through removal of biogenesis factors at critical checkpoints of 60S ribosome assembly (46, 47). In addition, previous studies found that midasin functions as a nuclear chaperone, is involved in the assembly/disassembly of macromolecular complexes in the nucleus, and is associated with maturation of 60S ribosome subunits (15). Our immuno histochemistry analysis showed that midasin was ubiquitously expressed in normal breast tissue as well as breast tumor tissue (**Figure 5**), suggesting its importance in normal physiologic functions. Interestingly, midasin levels were increased in ductal carcinoma compared to lobular carcinoma, but without additional tumor features such as molecular classification (hormone receptor status), tumor grade, node status, and tumor stage the exact relevance of this finding is yet to be understood. While the precise role of midasin in cancer has yet to be determined, the robust increase in protein expression observed in the mammospheres, the expression of midasin in ductal carcinoma tissue samples, along with its involvement in RFS, supports its involvement in both endocrine sensitive and endocrine resistant breast cancer.

Since ribosomal proteins play significant roles in normal cell functions, small and large ribosomal subunit deficiencies lead to distinct gene expression signatures that alter cell growth rates (48). This deregulation or altered expression is harmful to the

TABLE 3 | Normal breast tissue sample profile.

Patient Sample Identification	Sample type	Intensity	Subcellular localization	Fraction of cells stained	Age
Patient 2773	Adipocytes	weak	Cytoplasmic/membranous	<25%	23
	Glandular	moderate	Cytoplasmic/membranous, nuclear	>75%	23
	Myoepithelial	weak	Cytoplasmic/membranous	>75%	23
Patient 3544	Adipocytes	negative	Cytoplasmic/membranous	none	45
	Glandular	weak	Cytoplasmic/membranous Nuclear	25–75%	45
	Myoepithelial	weak	Cytoplasmic/membranous	25–75%	45

Normal breast tissue profile depicting sample type, staining intensity, subcellular localization, fraction of cells stained, and patient age.

TABLE 4 | Breast cancer patient tumor tissue profile.

Patient Sample Identification	Sample type	Staining Intensity	Subcellular localization	Fraction of cells stained	Age
Patient 2252	Lobular carcinoma	weak	Cytoplasmic/membranous	<25%	47
Patient 2898	Lobular carcinoma	weak	Cytoplasmic/membranous, nuclear	25–75%	41
Patient 3535	Lobular carcinoma	weak	Cytoplasmic/membranous	25–75%	47
Patient 2565	Lobular carcinoma	moderate	Cytoplasmic/membranous	>75%	51
Patient 2428	Ductal carcinoma	moderate	Cytoplasmic/membranous	>75%	75
Patient 1910	Ductal carcinoma	moderate	Cytoplasmic/membranous	>75%	61
Patient 2091	Ductal carcinoma	moderate	Cytoplasmic/membranous	>75%	40
Patient 3257	Ductal carcinoma	weak	Nuclear	25–75%	39

Tumor sample profile depicting sample type, staining intensity, subcellular localization, fraction of cells stained, and patient age.

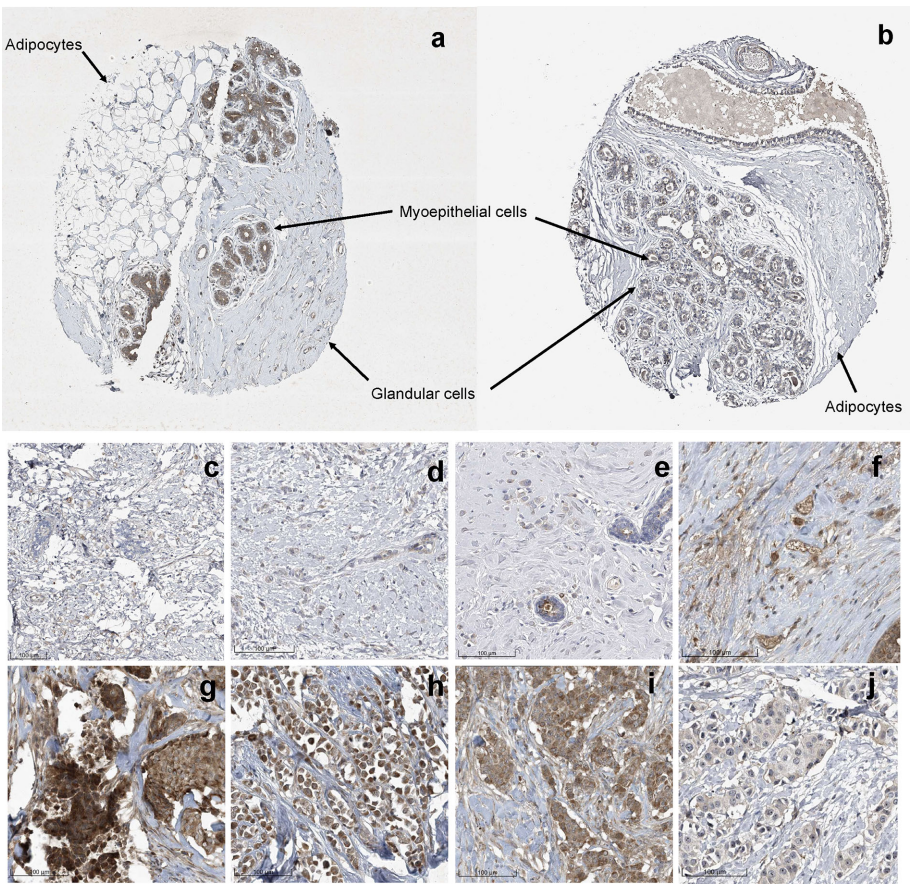


FIGURE 5 | Increased Midasin Expression Profile in Normal Breast and Breast Cancer Samples. Representative protein staining of midasin protein expression in normal mammary tissue **(A, B)**, lobular carcinoma tissue **(C–F)**, and ductal carcinoma tissue **(G–J)**. Image credit: Human Protein Atlas, www.proteinatlas.org, (Uhlén et al, 2015). Image available at the following URL: v19.proteinatlas.org/humancell.

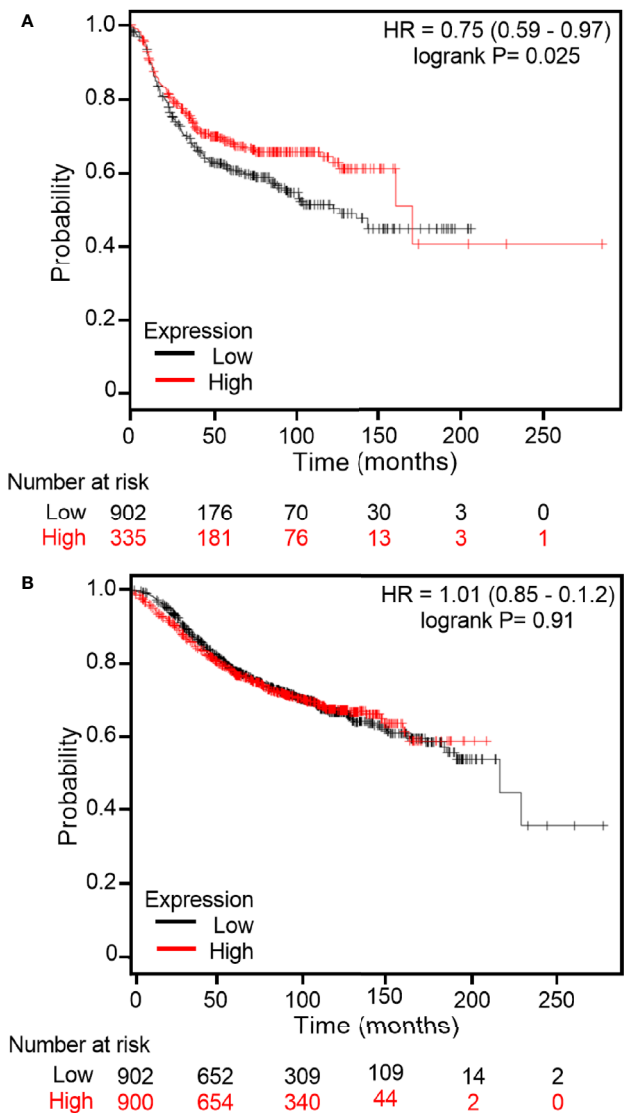


FIGURE 6 | Kaplan Meier plots of RFS based on midasin expression in ER+ and ER- breast cancer patients. Using Kaplan Meier Plotter, publicly available microarray repositories for breast cancer were interrogated to determine whether midasin expression was associated with different survival rates among **(A)** Estrogen Receptor negative (ER-) and **(B)** Estrogen Receptor positive (ER+) breast cancer patients. Hazard ratio (HR) and Logrank P values are shown. Low midasin expression (below median) is noted in black, and the high midasin expression (above median) is noted in red.

integrity and maintenance of the cell, causing serious consequences for cell fate. Some of the reported consequences on this ribosomal protein deregulation include defects in the synthesis of proteins relevant for tumorigenesis or cell survival, changes that affect the translation of specific mRNAs (ie., oncogenes and/or tumor suppressor genes), and alterations in the availability of ribosomes. To further investigate the protein expression profile of CSC-enriched letrozole resistant breast cancer cells, we conducted bioinformatics analyses. Our results show that many of the downregulated proteins were small ribosomal proteins associated with GO terms affiliated with the translational machinery (**Table 2**, **Figure 2**). Interestingly, when the down-regulated proteins were adjusted to control for false discovery rates (**Table S2**), the 40S small ribosomal proteins exhibited the greatest reduction in fold change compared to the 60S large ribosomal proteins. In a previous report, Fang and Zhang used integrated bioinformatics analyses and identified that low mRNA expression of RPL11, RPS14, RPS9, and RPL10A were related to a worse overall survival of breast cancer patients (49). This was notable since, in this study, we also observed decreased expression of RPL11 (-1.7-fold), RPS14 (-1.7-fold), and RPS9 (-2.5-fold) in the letrozole resistant mammospheres (**Figure 3**). While the role of RPS9 in breast cancer is unknown, decreased RPS9 expression has been observed in other solid tumors including pancreatic cancer (50).

It is well known that many cells transitioning from hormone dependent to hormone-refractory states no longer require estrogen for growth as they adapt and utilize MAPK and/or EGFR signaling pathways (6, 51). However, there may be other unknown regulatory mechanisms contributing to hormone independence. Here, we observed that midasin interacts with two downregulated proteins, EPRS and FARSA (**Figure 4**). Decreased expression of both proteins can affect transferring an amino acid from the tRNA onto a growing peptide and alter translation of specific proteins. While EPRS is necessary for the proliferation of tamoxifen-resistant ER+ breast cancer, its role in ER- breast cancer is unclear (37). Likewise, while a specific role for FARSA in cancer has not been identified, dysregulation of aminoacyl tRNA synthetases contribute to initiation, maintenance, and progression of carcinogenesis (52). Since the rate of cell growth has been demonstrated to be proportional to ribosomal biogenesis (53), it is likely that increased midasin and altered ribosomal protein expression could prime the translational machinery for increased protein synthesis or stabilize key proteins involved in cell proliferation.

It is well established that alterations in ribosomal proteins' expression could result in alterations in cellular homeostasis, affecting protein synthesis that could ultimately lead to cancer initiation or progression. To this end, it was critical to assess the impact of midasin expression levels on RFS in the ER+ and ER- breast cancer patients (**Figure 6**). There was no statistically significant difference in RFS within the ER+ cohort. However, high levels of midasin were more relevant in the ER- cohort and resulted in decreased RFS. For the first two years, the survival curves overlapped in the ER- cohort and initially, high levels of midasin seemed to be protective by extending survival up to an additional 15 years. However, beyond 15 years, high levels of midasin were associated with a worse prognosis among the ER- cohort, suggesting an early protective role for midasin within this population. It should be noted that although there was a trend in the association between high midasin levels and decreased RFS beyond 15 years, there were only 6 patients and a more robust sample size is needed to determine if this trend persists. While the precise role of midasin on survival is still not fully studied, its potential as a biomarker is very promising.

While the role of many of these AAA enzymes in breast cancer remains elusive, recent progress has revealed that midasin functions at successive maturation steps to remove ribosomal factors at critical transition points, first driving the exit of the early pre-60S particles from the nucleolus and then driving late pre-60S particles from the nucleus to the cytoplasm (46, 54). Ribosome biogenesis remains at the heart of translation thus requiring extensive regulation and coordination to meet the cellular demands of continuous ribosome production. As such, disruption of protein expression and interactions involved in ribosomal biogenesis affects translation, hence the overall fitness of the cell. This indicates that future studies are required to unravel a detailed mechanistic understanding of ribosome biogenesis which, in part, could aid in developing new strategies for targeting cancer stem cell enriched AI-resistant breast cancer.

DATA AVAILABILITY STATEMENT

The proteomic dataset is available at the Peptide Atlas repository through the link: https://db.systemsbiology.net/sbeams/cgi/PeptideAtlas/PASS_View?identifier=PASS01549.

AUTHOR CONTRIBUTIONS

KG and SL directed the project, conceived, and designed the experiments with supervision of ST. SL JP, RW, AD, WZ, and KZ contributed to data acquisition. WZ, KZ, and KG analyzed and interpreted the data. ST wrote the manuscript with support from KG. RW, JP, AD, and KG proofread the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was also supported in part by NIH grant SC1GM125617 awarded to ST. This publication was also made possible by funding from the Louisiana Cancer Research Consortium (LCRC) and the NIH-RCMI grant G12MD007595 from the National Institute on Minority Health and Health Disparities (NIMHD). The publication was also funded by NIH-RCMI grant U54MD007582 from the NIMHD. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the LCRC or the NIH. We also like to thank the Xavier University of Louisiana Cell and Molecular Biology Core Facility.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.540134/full#supplementary-material

Figure S1 | Western Blot analysis of midasin expression. Immunoblots of midasin protein expression in letrozole-sensitive (T47Darom), letrozole-resistant (T47DaromLR) breast cancer cells and MDA-MB 231 triple negative breast cancer cells cultured adherently (2D) or as mammospheres (3D). All cells were evaluated by immunoblot to examine the expression of midasin and GAPDH (loading control). **(A, C)** Graphs depict normalized percentages of protein expression intensities relative to 2D cell counterparts. **(B, D)** Representative immunoblot depicts the protein expression of midasin and GAPDH.

Table S1 | Proteomic analysis using BH method of upregulated proteins in LTLT-Ca mammospheres where the adjusted p value < 0.01.

Table S2 | Proteomic analysis using BH method of downregulated proteins in LTLT-Ca mammospheres where the adjusted p value < 0.01. Large ribosomal proteins are indicated by red bold text and small ribosomal proteins are indicated by blue bold text.

Table S3 | Gene Ontology analysis of the proteomic data of LTLT-Ca mammospheres.

Table S4 | KEGG Analysis of Proteomics data.

Table S5 | Letrozole Resistant Proteomic Analysis. Using a gel free proteomic approach combining TMT labeling, two-dimensional HPLC, and high-resolution mass spectrometry, over 1000 proteins were identified with quantitative abundance ratios present in both the LTLT-Ca mammospheres and the LTLT-Ca adherent breast cancer cells. Of the quantified proteins, 359 were significantly altered ($p < 0.05$), where 173 were upregulated (shown in yellow) and 186 downregulated shown in purple ($p < 0.05$, fold change >1.20).

REFERENCES

1. Johnston SR, Martin LA, Leary A, Head J, Dowsett M. Clinical strategies for rationale combinations of aromatase inhibitors with novel therapies for breast cancer. *J Steroid Biochem Mol Biol* (2007) 106(1-5):180–6. doi: 10.1016/j.jsbmb.2007.05.019
2. Miller WR, Larionov AA. Understanding the mechanisms of aromatase inhibitor resistance. *Breast Cancer Res* (2012) 14(1):201. doi: 10.1186/bcr2931
3. Brodie A, Jelovac D, Macedo L, Sabnis G, Tilghman S, Goloubeva O. Therapeutic observations in MCF-7 aromatase xenografts. *Clin Cancer Res* (2005a) 11(2 Pt 2):884s–8s.
4. Brodie A, Jelovac D, Sabnis G, Long B, Macedo L, Goloubeva O. Model systems: mechanisms involved in the loss of sensitivity to letrozole. *J Steroid Biochem Mol Biol* (2005b) 95(1-5):41–8. doi: 10.1016/j.jsbmb.2005.04.026
5. Sabnis G, Schayowitz A, Goloubeva O, Macedo L, Brodie A. Trastuzumab reverses letrozole resistance and amplifies the sensitivity of breast cancer cells to estrogen. *Cancer Res* (2009) 69(4):1416–28. doi: 10.1158/0008-5472.can-08-0857
6. Tilghman SL, Townley I, Zhong Q, Carriere PP, Zou J, Llopis SD, et al. Proteomic signatures of acquired letrozole resistance in breast cancer: suppressed estrogen signaling and increased cell motility and invasiveness. *Mol Cell Proteomics* (2013) 12(9):2440–55. doi: 10.1074/mcp.M112.023861
7. Carriere PP, Llopis SD, Naiki AC, Nguyen G, Phan T, Nguyen MM, et al. Glyceollin I Reverses Epithelial to Mesenchymal Transition in Letrozole Resistant Breast Cancer through ZEB1. *Int J Environ Res Public Health* (2016) 13(1):ijerph13010010. doi: 10.3390/ijerph13010010
8. Gilani RA, Kazi AA, Shah P, Schech AJ, Chumsri S, Sabnis G, et al. The importance of HER2 signaling in the tumor-initiating cell population in aromatase inhibitor-resistant breast cancer. *Breast Cancer Res Treat* (2012) 135(3):681–92. doi: 10.1007/s10549-012-2148-8
9. Chang JC. Cancer stem cells: Role in tumor growth, recurrence, metastasis, and treatment resistance. *Medicine* (2016) 95(1 Suppl 1):S20–5. doi: 10.1097/MD.0000000000004766

10. Silvera D, Formenti SC, Schneider RJ. Translational control in cancer. *Nat Rev Cancer* (2010) 10(4):254–66. doi: 10.1038/nrc2824
11. Fromont-Racine M, Senger B, Saveanu C, Fasiolo F. Ribosome assembly in eukaryotes. *Gene* (2003) 313:17–42. doi: 10.1016/s0378-1119(03)00629-2
12. Tschochner H, Hurt E. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol* (2003) 13(5):255–63. doi: 10.1016/s0962-8924(03)00054-0
13. Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mougin A, et al. The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* (2008) 65(15):2334–59. doi: 10.1007/s00018-008-8027-0
14. Woolford JL Jr, Baserga SJ. Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics* (2013) 195(3):643–81. doi: 10.1534/genetics.113.153197
15. Garbarino JE, Gibbons IR. Expression and genomic analysis of midasin, a novel and highly conserved AAA protein distantly related to dynein. *BMC Genomics* (2002) 3:18. doi: 10.1186/1471-2164-3-18
16. Iyer LM, Leipe DD, Koonin EV, Aravind L. Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* (2004) 146(1-2):11–31. doi: 10.1016/j.jsb.2003.10.010
17. Kawashima SA, Chen Z, Aoi Y, Patgiri A, Kobayashi Y, Nurse P, et al. Potent, Reversible, and Specific Chemical Inhibitors of Eukaryotic Ribosome Biogenesis. *Cell* (2016) 167(2):512–524.e514. doi: 10.1016/j.cell.2016.08.070
18. Romes EM, Sobhany M, Stanley RE. The Crystal Structure of the Ubiquitin-like Domain of Ribosome Assembly Factor Ytm1 and Characterization of Its Interaction with the AAA-ATPase Midasin. *J Biol Chem* (2016) 291(2):882–93. doi: 10.1074/jbc.M115.693259
19. Galani K, Nissan TA, Petfalski E, Tollervey D, Hurt E. Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits. *J Biol Chem* (2004) 279(53):55411–8. doi: 10.1074/jbc.M406876200
20. Nissan TA, Galani K, Maco B, Tollervey D, Aebi U, Hurt E. A pre-ribosome with a tadpole-like structure functions in ATP-dependent maturation of 60S subunits. *Mol Cell* (2004) 15(2):295–301. doi: 10.1016/j.molcel.2004.06.033
21. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? *Nat Rev Cancer* (2003) 3(3):179–92. doi: 10.1038/nrc1015
22. Sulima SO, Hofman IJF, De Keersmaecker K, Dinman JD. How Ribosomes Translate Cancer. *Cancer Discovery* (2017) 7(10):1069. doi: 10.1158/2159-8290.CD-17-0550
23. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* (2000) 100(1):57–70. doi: 10.1016/s0092-8674(00)81683-9
24. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
25. Ellis SR, Gleizes PE. Diamond Blackfan anemia: ribosomal proteins going rogue. *Semin Hematol* (2011) 48(2):89–96. doi: 10.1053/j.seminhematol.2011.02.005
26. Ajore R, Raiser D, McConkey M, Joud M, Boidol B, Mar B, et al. Deletion of ribosomal protein genes is a common vulnerability in human cancer, especially in concert with TP53 mutations. *EMBO Mol Med* (2017) 9(4):498–507. doi: 10.15252/emmm.201606660
27. Chiew GGY, Wei N, Sultania S, Lim S, Luo KQ. Bioengineered three-dimensional co-culture of cancer cells and endothelial cells: A model system for dual analysis of tumor growth and angiogenesis. *Biotechnol Bioeng* (2017) 114(8):1865–77. doi: 10.1002/bit.26297
28. Long BJ, Jelovac D, Handratta V, Thiantanawat A, MacPherson N, Ragaz J, et al. Therapeutic strategies using the aromatase inhibitor letrozole and tamoxifen in a breast cancer model. *J Natl Cancer Institute* (2004) 96(6):456–65. doi: 10.1093/jnci/djh076
29. Gupta A, Mehta R, Alimirah F, Peng X, Murillo G, Wiehle R, et al. Efficacy and mechanism of action of Proellex, an antiprogestin in aromatase overexpressing and Letrozole resistant T47D breast cancer cells. *J Steroid Biochem Mol Biol* (2013) 133:30–42. doi: 10.1016/j.jsbmb.2012.08.004
30. Rhodes LV, Tilghman SL, Boue SM, Wang S, Khalili H, Muir SE, et al. Glyceollins as novel targeted therapeutic for the treatment of triple-negative breast cancer. *Oncol Lett* (2012) 3(1):163–71. doi: 10.3892/ol.2011.460
31. Capes-Davis A, Reid YA, Kline MC, Storts DR, Strauss E, Dirks WG, et al. Match criteria for human cell line authentication: where do we draw the line? *Int J Cancer* (2013) 132(11):2510–9. doi: 10.1002/ijc.27931
32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* (2009) 4(1):44–57. doi: 10.1038/nprot.2008.211
33. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* (2010) 123(3):725–31. doi: 10.1007/s10549-009-0674-9
34. Gyorffy B, Lanczky A, Szallasi Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocr Relat Cancer* (2012) 19(2):197–208. doi: 10.1530/erc-11-0329
35. Johnson AW. Ribosomes: lifting the nuclear export ban. *Curr Biol* (2014) 24(3):R127–129. doi: 10.1016/j.cub.2013.12.013
36. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Society: Ser B (Methodological)* (1995) 57(1):289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
37. Karunarathna U, Kongsema M, Zona S, Gong C, Cabrera E, Gomes AR, et al. OTUB1 inhibits the ubiquitination and degradation of FOXM1 in breast cancer and epirubicin resistance. *Oncogene* (2016) 35(11):1433–44. doi: 10.1038/onc.2015.208
38. Katsyv I, Wang M, Song WM, Zhou X, Zhao Y, Park S, et al. EPRS is a critical regulator of cell proliferation and estrogen signaling in ER+ breast cancer. *Oncotarget* (2016) 7(43):69592–605. doi: 10.18632/oncotarget.11870
39. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science* (2015) 347(6220):1260419. doi: 10.1126/science.1260419
40. Tilghman SL, Sabnis G, Brodie AM. Upregulation of AIB1, aromatase and ERalpha provides long-term estrogen-deprived human breast cancer cells with a mechanistic growth advantage for survival. *Horm Mol Biol Clin Investig* (2011) 3(1):357–66. doi: 10.1515/hmbci.2010.042
41. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ, et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* (2003) 17(10):1253–70. doi: 10.1101/gad.1061803
42. Morrison BJ, Hastie ML, Grewal YS, Bruce ZC, Schmidt C, Reynolds BA, et al. Proteomic comparison of mcf-7 tumoursphere and monolayer cultures. *PloS One* (2012) 7(12):e52692. doi: 10.1371/journal.pone.0052692
43. Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D, et al. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res* (2005) 65(13):5506–11. doi: 10.1158/0008-5472.can-05-0626
44. Cariati M, Naderi A, Brown JP, Smalley MJ, Pinder SE, Caldas C, et al. Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int J Cancer* (2008) 122(2):298–304. doi: 10.1002/ijc.23103
45. Ao A, Morrison BJ, Wang H, Lopez JA, Reynolds BA, Lu J. Response of estrogen receptor-positive breast cancer tumorspheres to antiestrogen treatments. *PloS One* (2011) 6(4):e18810. doi: 10.1371/journal.pone.0018810
46. Ulbrich C, Diepholz M, Bassler J, Kressler D, Pertschy B, Galani K, et al. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell* (2009) 138(5):911–22. doi: 10.1016/j.cell.2009.06.045
47. Bassler J, Kallas M, Pertschy B, Ulbrich C, Thoms M, Hurt E. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Mol Cell* (2010) 38(5):712–21. doi: 10.1016/j.molcel.2010.05.024
48. Cheng Z, Mugler CF, Keskin A, Hodapp S, Chan LY, Weis K, et al. Small and Large Ribosomal Subunit Deficiencies Lead to Distinct Gene Expression Signatures that Reflect Cellular Growth Rate. *Mol Cell* (2019) 73(1):36–47.e10. doi: 10.1016/j.molcel.2018.10.032
49. Fang E, Zhang X. Identification of breast cancer hub genes and analysis of prognostic values using integrated bioinformatics analysis. *Cancer Biomark* (2017) 21(1):373–81. doi: 10.3233/cbm-170550
50. Crnogorac-Jurcevic T, Efthimiou E, Capelli P, Blaveri E, Baron A, Terris B, et al. Gene expression profiles of pancreatic cancer and stromal desmoplasia. *Oncogene* (2001) 20(50):7437–46. doi: 10.1038/sj.onc.1204935
51. Jelovac D, Sabnis G, Long BJ, Macedo L, Goloubeva OG, Brodie AM. Activation of mitogen-activated protein kinase in xenografts and cells

during prolonged treatment with aromatase inhibitor letrozole. *Cancer Res* (2005) 65(12):5380–9. doi: 10.1158/0008-5472.can-04-4502
52. Kim D, Kwon NH, Kim S. Association of aminoacyl-tRNA synthetases with cancer. *Top Curr Chem* (2014) 344:207–45. doi: 10.1007/128_2013_455
53. Montanaro L, Trere D, Derenzini M. Nucleolus, ribosomes, and cancer. *Am J Pathol* (2008) 173(2):301–10. doi: 10.2353/ajpath.2008.070752
54. Kressler D, Hurt E, Bergler H, Bassler J. The power of AAA-ATPases on the road of pre-60S ribosome maturation–molecular machines that strip pre-ribosomal particles. *Biochim Biophys Acta* (2012) 1823(1):92–100. doi: 10.1016/j.bbamcr.2011.06.017

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 Gallegos, Patel, Llopis, Walker, Davidson, Zhang, Zhang and Tilghman. 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 Oncology

ORIGINAL RESEARCH
published: 09 March 2021
doi: 10.3389/fonc.2021.602712

Protein and Signaling Pathway Responses to rhIL-6 Intervention Before Lobaplatin Treatment in Osteosarcoma Cells

Huan Wang[1†], *Bin Li*[1†], *Kang Yan*[1], *Yonghong Wu*[1], *Yanhua Wen*[1], *Yunyan Liu*[1], *Pei Fan*[2*] *and Qiong Ma*[1*]

[1] Orthopedic Oncology Institute, Department of Orthopedic Surgery, Tangdu Hospital, Fourth Military Medical University, Xi'an, China, [2] Department of Orthopedics, The Second Affiliated Hospital of Wenzhou Medical University, Yuying Children's Hospital, Wenzhou, China

Lobaplatin is a third-generation platinum-based antineoplastic agent and is widely used for osteosarcoma treatment before and after tumor removal. However, treatment failure often results from lobaplatin drug resistance. In our study, we found that SaOS-2 and SOSP-9607 osteosarcoma cells became less sensitive to lobaplatin after treatment with exogenous interleukin (IL)-6. Quantitative proteomic analysis was performed to elucidate the underlying mechanism in SaOS-2 osteosarcoma cells. Cells were divided into a control group (CG), a lobaplatin treatment group (LG), a recombinant human IL-6 (rhIL-6), and a lobaplatin treatment group (rhILG). We performed three biological replicates in each group to compare the differential protein expression between groups using a tandem mass tag (TMT) labeling technology based on liquid chromatography-tandem mass spectrometry (LC-MS/MS). A total of 1,313 proteins with significant differential expression was identified and quantified. The general characteristics of the significantly enriched proteins were identified by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, and protein–protein interaction (PPI) analysis was conducted using IntAct and STRING. In total, 31 proteins were further verified by parallel reaction monitoring (PRM), among which ras GTPase-activating protein-binding protein 1 (G3BP1), fragile X mental retardation syndrome-related protein 1 (hFXR1p), and far upstream element-binding protein 1 (FUBP1) were significantly differentially expressed. Immunohistochemistry results showed that these three proteins are highly expressed in specimens from platinum-resistant osteosarcoma patients, while the proteins are negatively or weakly expressed in specimens from platinum-sensitive osteosarcoma patients. The immunofluorescence staining results were in accord with the immunohistochemistry staining results. siRNA knockdown of FUBP1 showed a strikingly decreased IC50 value for lobaplatin in FUBP1-silenced cells, which verified the role of FUBP1 in the drug susceptibility of osteosarcoma and the potential therapeutic value for increasing the sensitivity to lobaplatin. This is the first proteomic study on a rhIL-6 intervention before lobaplatin treatment in osteosarcoma cells.

Keywords: osteosarcoma, rhIL-6 intervention, lobaplatin, liquid chromatography-tandem mass spectrometry, protein–protein interaction, FUBP1

OPEN ACCESS

Edited by:
Bhaswati Chatterjee, National Institute of Pharmaceutical Education and Research, India

Reviewed by:
Gongping Sun, Shandong University, China
Gaurav Kandoi, Iowa State University, United States

***Correspondence:**
Qiong Ma
maqiong@fmmu.edu.cn
Pei Fan
fanpei@wmu.edu.cn

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *04 September 2020*
Accepted: *16 February 2021*
Published: *09 March 2021*

Citation:
Wang H, Li B, Yan K, Wu Y, Wen Y, Liu Y, Fan P and Ma Q (2021) Protein and Signaling Pathway Responses to rhIL-6 Intervention Before Lobaplatin Treatment in Osteosarcoma Cells. Front. Oncol. 11:602712. doi: 10.3389/fonc.2021.602712

INTRODUCTION

Osteosarcoma accounts for approximately 44.6% of malignant bone tumors and is the most common primary malignant bone tumor clinically. Osteosarcoma usually occurs in children and young adolescents and demonstrates poor prognosis (1, 2). As osteosarcoma is considered less sensitive to radiotherapy, chemotherapy has been the traditional treatment accompanying surgery to improve the survival of tumor-suffering patients (3). Platinum-based drugs, such as cisplatin, carboplatin, and lobaplatin, are extensively used for osteosarcoma patients (4, 5). Lobaplatin is a third-generation platinum-based chemotherapy drug with ideal antitumor activity and stability. However, osteosarcoma becomes less sensitive to lobaplatin after one or two courses of treatment.

Interleukin (IL)-6 is a pleiotropic cytokine that can act as a resistance factor against some antitumor drugs, such as doxorubicin, VP-16, and cisplatin (6). It has been reported that elevated IL-6 is associated with neoplastic growth and apoptosis (7–9). However, the underlying comprehensive mechanism of IL-6 treatment on the decreased sensitivity to platinum-based chemotherapeutics in osteosarcoma cells has not yet been systemically elucidated.

Proteomic analysis has shown great strength for large-scale protein investigations to explain the specific mechanism of the role of exogenous IL-6 in drug treatment for osteosarcoma. However, there are no studies on the proteomic response to IL-6 usage before lobaplatin treatment at the protein level. In this study, we investigated the differential protein expression after recombinant human IL-6 (rhIL-6) intervention before lobaplatin treatment of an osteosarcoma SaOS-2 cell model and revealed potential biomarkers that indicate the decreased sensitivity of osteosarcoma to lobaplatin.

MATERIALS AND METHODS

Cell Culture and Treatment

The human osteosarcoma cell lines SaOS-2 and SOSP-9607 were maintained in our laboratory (10) and grown in Dulbecco's modified Eagle's medium and RPMI 1640 medium (HyClone, USA) with 10% fetal bovine serum (Gibco, USA). Cells were cultured at 37°C in a humidified atmosphere with 5% CO_2. Recombinant human IL-6 (PeproTech, USA) was used to pretreat tumor cells before lobaplatin treatment.

Cell Viability Assay

Cells were seeded in a 96-well plate at 6×10^3 cells per well and divided into the following four groups: CG (cells treated with the solvent of rhIL-6 and the solvent of lobaplatin), rhIL-6G (cells treated with 60 ng/mL rhIL-6 and the solvent of lobaplatin), LG (cells treated with the solvent of rhIL-6 and 10 μg/mL lobaplatin 8 h later), and rhILG (cells treated with 60 ng/mL rhIL-6 and 10 μg/mL lobaplatin 8 h later). Each group had three replicates. CCK-8 assay was then performed after 24 h of lobaplatin treatment according to the manufacturer's instructions (Dojindo, Japan), and the absorbance at 450 nm was measured to detect the cell viability with a multiscan reader (Thermo Scientific, USA). The results are expressed as the mean optical density ± SEM ($n = 3$).

Apoptosis Assessment

Apoptotic cells were measured using an annexin V-FITC/PI detection kit (BD Biosciences, USA) and analyzed by flow cytometry (Beckman Coulter, USA) according to the manufacturer's instructions. A total of 3×10^5 cells were seeded in a 6-well plate, the treatment was the same as described previously, and each group contained three replicates. Cells were then collected by gentle digestion with trypsin, washed with precooled PBS, and resuspended in 400 μL of binding buffer. Five microliters of annexin V-FITC and 1 μg/mL PI were used to stain cells in the CG, rhIL-6G, LG, and rhILG for 15 min in the dark and then analyzed immediately. The early apoptotic cells were annexin V+/PI–, and the late apoptotic cells were annexin V+/PI+.

The cell nuclei were stained with Hoechst 33342 solution (Roche, Switzerland). Cells were seeded in a 24-well plate at 7×10^4 cells per well and divided into the CG, rhIL-6G, LG, and rhILG. The treatment was the same as described previously, and each group contained three replicates. The cultures were stained in the incubator for 15 min at 37°C at a concentration of 5 μmol/ml. After incubation, the cells were washed with PBS three times, and the fluorescence was detected with a fluorescence microscope (Olympus, Japan).

Proteomics Analysis

Sample Preparation

A total of 8×10^6 cells were seeded in nine 15-cm culture dishes, and each group contained three dishes for replicates. After the treatment as previously described, cells were washed three times and collected by scraping in precooled PBS on ice. Cell pellets were centrifuged at 4°C, and the supernatants were removed. Samples were then rapidly frozen in liquid nitrogen for 30 s. Cell pellets were lysed with SDT buffer (4% SDS and 100 mM Tris-HCl at pH 7.6), boiled for 15 min and centrifuged at 14,000 × g for 40 min. The supernatants were collected and quantified with the BCA Protein Assay Kit (Pierce, USA). Samples were stored at −80°C until use. A total of 20 μg of proteins for each sample was mixed with 5X loading buffer and boiled for 5 min. The proteins were separated on a 12.5% SDS-PAGE gel. Protein bands were visualized by Coomassie Blue R-250 staining.

For the filter-aided sample preparation (FASP) and digestion (11), 200 μg of proteins from each sample was incorporated into 30 μL SDT buffer (4% SDS, 100 mM DTT, and 150 mM Tris-HCl at pH 8.0). DTT and other low-molecular-weight components were removed with uric acid (UA) buffer (8 M urea and 150 mM Tris-HCl at pH 8.5) by repeated ultrafiltration (Sartorius, 30 kDa). Then, 100 μL of 100 mM iodoacetamide (IAA) in UA buffer was added to block reduced cysteine residues, and the samples were incubated for 30 min under darkness. The filters were washed three times with 100 μL UA buffer and then twice with 100 μL of 0.1 M tetraethylammonium bromide (TEAB) buffer. Finally, the protein suspensions were digested with 4 μg trypsin (Promega, USA) in 40 μL 0.1 M TEAB buffer overnight at 37°C, and the

resulting peptides were collected. By calculating the frequency of tryptophan and tyrosine in vertebrate proteins, an extinction coefficient of 1.1 of 0.1% (g/L) solution was chosen to estimate the peptide concentration using a UV spectrometer at 280 nm.

A 100 μg peptide mixture of each sample was labeled using the tandem mass tag (TMT) reagent according to the manufacturer's instructions (Thermo Fisher Scientific, USA). The peptides in the CG, LG, and rhILG were labeled with TMT-126, TMT-127N, TMT-127C, TMT-128N, TMT-128C, TMT-129N, TMT-129C, TMT-130C, and TMT-131, with three biological replicates. TMT-labeled peptides were fractionated by reversed-phase (RP) chromatography using an Agilent 1260 Infinity II HPLC. The peptide mixture was diluted with buffer A [10 mM $HCOONH_4$ and 5% acetonitrile (ACN) at pH 10.0] and loaded onto an XBridge Peptide BEH C18 Column, 130 Å, 5 μm, 4.6 × 100 mm column. The peptides were eluted at a flow rate of 1 ml/min with a gradient of 0–7% buffer B (10 mM $HCOONH_4$ and 85% ACN at pH 10.0) for 5 min, 7–40% buffer B for 5–40 min, 40–100% buffer B for 45–50 min, and 100% buffer B for 50–65 min. The elution was monitored by the absorbance at 214 nm, and fractions were collected every 1 min for 50 min. The collected fractions were dried down via vacuum centrifugation.

MS Analysis

Each fraction was injected for nanoliquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) analysis. The peptide mixture was loaded onto a C18 RP analytical column (Thermo Fisher Scientific, Acclaim PepMap RSLC 50 μm × 15 cm, nanoViper, P/N164943) in buffer A (0.1% formic acid) and separated with a 1.5-h linear gradient of buffer B (80% ACN and 0.1% formic acid) at a flow rate of 300 nl/min. Specifically, the peptides were eluted as follows: 6% buffer B for 5 min, 6–28% buffer B for 63 min, 28–38% buffer B for 10 min, 38–100% buffer B for 7 min, and 100% buffer B for 5 min. Liquid chromatography-tandem mass spectrometry analysis was performed on a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, USA) that was coupled to an Easy nLC (Thermo Fisher Scientific, USA) for 90 min. The mass spectrometer was operated in the positive ion mode. MS data were acquired using a data-dependent top 10 method dynamically choosing the most abundant precursor ions from the survey scan (350–1,800 m/z) for higher energy collisional dissociation (HCD) fragmentation. The automatic gain control (AGC) target was set to 3.0E6, and the maximum injection time was 45 ms. Survey scans were acquired at a resolution of 70,000 at 200 m/z, the resolution for the HCD spectra was set to 17,500 at 200 m/z, and the isolation width was 2 m/z. The normalized collision energy (NCE) was 30 eV.

MS/MS spectra were searched using the MASCOT engine (Matrix Science, UK; version 2.6, RRID: SCR_014322) embedded in Proteome Discoverer 2.1. The principal parameters were set as follows: peptide false discovery rate (FDR) ≤ 0.01 and Fragment Mass Tolerance of 0.1 Da. Proteins with a fold change>1.2 or <0.83 and $p < 0.05$ (Student's t-test) were considered to be differentially expressed proteins.

GO and KEGG Pathway Analyses

All protein sequences were aligned to the Homo sapiens database that was downloaded from NCBI (ncbi-blast-2.2.28+-win32.exe), and only the sequences in the top 10 with an E-value < 1E-3 were retained. Then, the GO terms (database version: go_201910.obo, RRID: SCR_002811) of the sequences with the top bit score were selected by Blast2GO, and the annotations of GO terms to proteins were completed by the Blast2GO Command Line. After the elementary annotation, InterProScan was used to search the EBI database by motif, and the functional information for each motif was added to the proteins to improve the annotation. Then, further improvements to the annotation and connections between GO terms were carried out by ANNEX. Fisher's exact test was used to enrich GO terms by comparing the number of differentially expressed proteins and total proteins correlated to the GO terms. Pathway analysis was performed using the KEGG database (database version: KO_INFO_END_20191021, RRID: SCR_012773). Fisher's exact test was used to identify the significantly enriched pathways by comparing the number of differentially expressed proteins and total proteins correlated to pathways.

Cluster Analysis

Cluster analysis was performed using STEM (12) to ensure the changing trends of various kinds of proteins and proteins with similarly changing trends were classified in a cluster accordingly. Significance analysis was performed for each cluster, and clusters with statistical significance were acquired.

PPI Analysis

The IntAct molecular interaction database (database version: IntAct View 4.2.16, SCR_006944) was utilized to study the relationship between the differentially expressed proteins using their distinct gene symbols, and the Cytoscape software was used to visualize the functional PPI networks. Protein–protein interaction (PPI) information was also retrieved from the STRING database, which shows direct experimental interactions and predicted interactions using computational algorithms. In both of the above databases, the degrees of connectivity for each differentially expressed protein between the LG and rhILG were calculated to evaluate the importance of proteins in the PPI network. The intersection proteins of the two databases were considered to be the target proteins.

PRM Validation

Parallel reaction monitoring (PRM) validation was performed among the differentially expressed proteins to verify the proteomic analysis based on the TMT label-based LC-MS/MS. In total, 31 proteins were analyzed, including three proteins of interest, namely, ras GTPase-activating protein-binding protein 1 (G3BP1), far upstream element-binding protein 1 (FUBP1), and fragile X mental retardation syndrome-related protein 1 (hFXR1p). Two micrograms of peptide mixture was loaded onto the C18 RP analytical column (Thermo Fisher Scientific, Acclaim PepMap RSLC 50 μm ×15 cm, nanoViper, P/N164943) in buffer A (0.1% formic acid) and separated with a non-linear gradient of buffer B (80% ACN and 0.1% formic acid) at a flow rate of 300

nl/min. The gradient was 2–8% buffer B for 1 min, 8–28% buffer B for 45 min, 28–40% buffer B for 50 min, 40–90% buffer B for 56 min, and 90% buffer B for 3 min.

Peptide fragmentation and targeted PRM MS were performed using a Q Exactive Mass Spectrometer (Thermo Scientific, USA). The mass spectrometer was operated in the positive ion mode. MS data were acquired with the following settings: (1) Full-MS: scan range (m/z) = 350–1,500; resolution = 60,000; AGC target = 1E6; and maximum injection time = 50 ms; (2) PRM: resolution = 17,500; AGC target = 1E5; maximum injection time = 50 ms; isolation window = 2 m/z; and NCE = 27%. The MS RAW file was converted to the mzXML format via the MSConvertGUI software. The resulting mass spectrum mzXML file was analyzed using the Skyline 3.6 software for PRM data.

Proteome Discoverer (v. 2.1) and Skyline (v. 3.6) were used for raw MS data processing. The MASCOT engine was applied to search the derived peak list according to the UniProt Homo sapiens protein database in Proteome Discoverer. The peptides were generated using trypsin as the enzyme. The precursor mass tolerance was specified as 10 ppm, and for MS2 fragments, the tolerance was 0.05 Da. Carbamidomethyl was chosen as the fixed modification, while acetyl and oxidation were set as the variable modifications. A reverse database search strategy was used with the peptide and protein FDR set to 1%. The default settings were applied to all other operations. A spectral library with a cut-off score of 0.99 was built with the MS/MS table file output by MASCOT in Skyline.

Clinical Specimens, Immunohistochemical, and Immunofluorescence Staining

Specimens were collected from 30 osteosarcoma patients who underwent platinum-based chemotherapy and tumor resection at Tangdu Hospital, Fourth Military Medical University between January 2008 and December 2018. Tumors with regression of more than 50% were categorized into the chemotherapy-sensitive group according to a previous study that evaluated the chemotherapy response in rectal cancer (13). In total, 15 patients were categorized into the chemotherapy-sensitive group, and the rest were categorized into the chemotherapy-resistant group. Studies concerning patients and their specimens were approved by the Ethics Committee of the Fourth Military Medical University, and written informed consent was provided. Formalin-fixed paraffin-embedded specimens were then sectioned into 4–6-μm sections (LEICA, Germany), affixed onto microscope slides, and heated in an oven at 60°C overnight. The following procedures were performed according to the instructions provided by Abcam Company (Abcam, USA). The anti-G3BP (Abcam, Cat# ab56574, RRID: AB_941699), anti-FXR1 (Abcam, Cat#ab129089, RRID: AB_11154960), and anti-FUBP1 (Abcam, Cat # ab 181111) antibodies were used at dilutions of 1:100, 1:100, and 1:250, respectively. DAB (GK347010, Gene Tech) or fluorescent secondary antibody (Affinity Biosciences, Cat# S0011, RRID: AB_2844800) was used to detect proteins expression. Two different pathologists independently scored the tissues, which were assigned based on the positive cell numbers combined with the intensity of staining as described previously (14).

siRNAs and Transfection

Small interfering RNA (siRNA) oligonucleotides sequences were synthesized (GenePharma, China). The siRNA sequences targeting human G3BP1 are: RNAi#1, 5′-GGGAAUUUGU GAGACAGUAT T-3′ (sense), 5′-UACUGUCUCACAAAU UCCCTT-3′ (antisense); RNAi#2, 5′-GCCUGAGCCAGUAUUA GAATT-3′ (sense), 5′-U UCU AAUACUGGCUCAGGCTT-3′ (antisense); and RNAi#3, 5′-GCGAGAACAACGAAUAAAUTT-3′ (sense), 5′-AUUUAUUCGUUGUUCUCGCTT-3′ (antisense). The siRNA sequences targeting human FUBP1 are: RNAi#1, 5′-GGUGUUCGC AUUCAGUUUATT-3′ (sense), 5′-UAAAC UGAAUGCGAACACCTT-3′ (antisense); RNAi#2, 5′-GGUGC UGACAAACCUCUUATT-3′ (sense), 5′ -UAA GAGGU UUGUCAGCACCTT-3′ (antisense); and RNAi#3, 5′-CGGCA ACUCAUAGAAGAAATT-3′ (sense), 5′-UUUCUUCUAUG AGUUGCCGTT-3′ (antisense). The siRNA sequences targeting human FXR1 are: RNAi#1, 5′-GGAGC UGACGGUGGAG GUUTT-3′ (sense), 5′-AACCUCCACCGUCAGCUCCTT-3′ (antisense); RNAi#2, 5′-GCAAAUGACCAAGAGCCAUTT-3′ (sense), 5′-AUGGCUCUUGGUCAUUUGCTT-3′ (antisense); and RNAi #3,5′-GCU AGAGGUUUCUUGGAAUTT-3′ (sense), 5′-AUUCCAAGAAACCUCUAGCTT-3′ (antisense). A scrambled siRNA was used as the control. siRNA transfection was performed using Lipofectamine 3000 (Invitrogen, USA) according to the manufacturer's instruction.

Statistical Analysis

Each experiment was carried out at least three times, and the mean value was calculated. Statistical analysis was performed with SPSS 25.0 (IBM, RRID: SCR_002865) and GraphPad Prism 7 software (GraphPad, RRID: SCR_002798). Data are presented as the mean ± SEM. The IC50 was calculated by regression analysis. Cell viability was obtained using one-way ANOVA. Differences between two groups were analyzed by Student's *t*-test, and $p < 0.05$ was considered statistically significant.

RESULTS

Exogenous IL-6 Reduces the Susceptibility of SaOS-2 and SOSP-9607 Osteosarcoma Cells to Lobaplatin

To assess the potential role of IL-6 in the process of lobaplatin resistance in osteosarcoma cells, 60 ng/mL rhIL-6 was applied to SaOS-2 and SOSP-9607 osteosarcoma cells 8 h before lobaplatin treatment. After 24 h, a Cell Counting Kit-8 (CCK-8) assay was performed to detect the viability of tumor cells. It was demonstrated that the average cell viability of SaOS-2 and SOSP-9607 osteosarcoma cells treated with rhIL-6 and then with lobaplatin (rhIL-6 and lobaplatin group, rhILG) was 73.58% ± 0.021 and 40.86% ± 0.054, while that of SaOS-2 and SOSP-9607 cells treated only with lobaplatin (lobaplatin group, LG) was 52.65% ± 0.0009 and 27.66% ± 0.038, respectively, indicating that SaOS-2 and SOSP-9607 cells pretreated with rhIL-6 exhibit greater resistance to lobaplatin ($p = 0.0093$

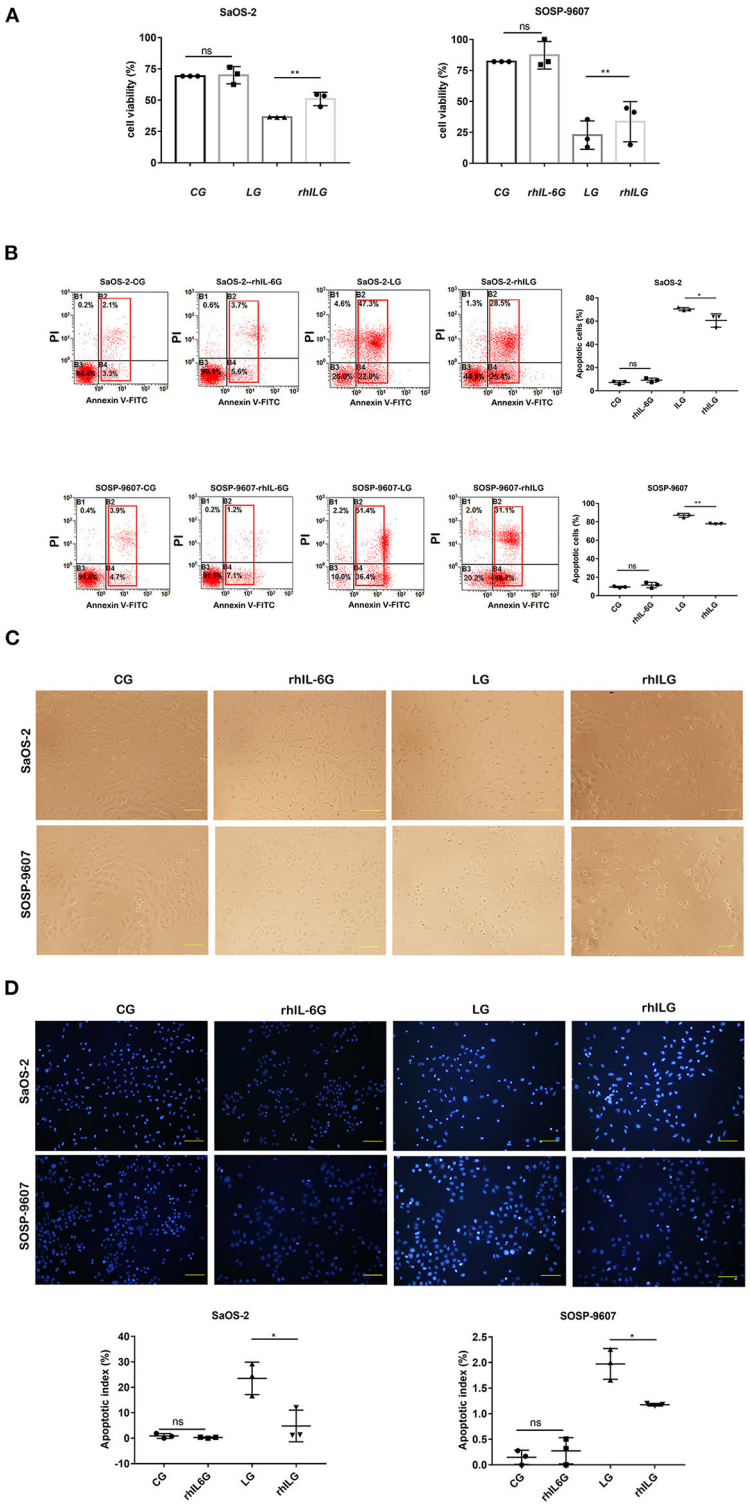


FIGURE 1 | rhIL6 intervention before lobaplatin treatment of osteosarcoma cells increased cell viability and reduced SaOS-2 and SOSP-9607 cell apoptosis. **(A)** Cell viability measured by CCK-8 assay in the CG, rhIL-6G, LG, and rhILG. Experiments were performed with at least three biological replicates, and a $p \leq 0.05$ was considered statistically significant. **(B)** Cells in each group were stained with annexin V and PI and measured by flow cytometry. Early and late apoptotic cells were demonstrated as annexin V+/PI– (early apoptosis) and annexin V+/PI+ (late apoptosis), respectively. **(C)** Morphological differences in osteosarcoma cells in the CG, rhIL-6G, LG, and rhILG were observed by bright-field microscopy (×200), scale bar = 100 μm. **(D)** Morphological differences in osteosarcoma cell nuclei were detected by Hoechst 33342 staining and observed by fluorescence microscopy (×200), scale bar = 100 μm. $^{*}p < 0.05$; $^{**}p < 0.01$.

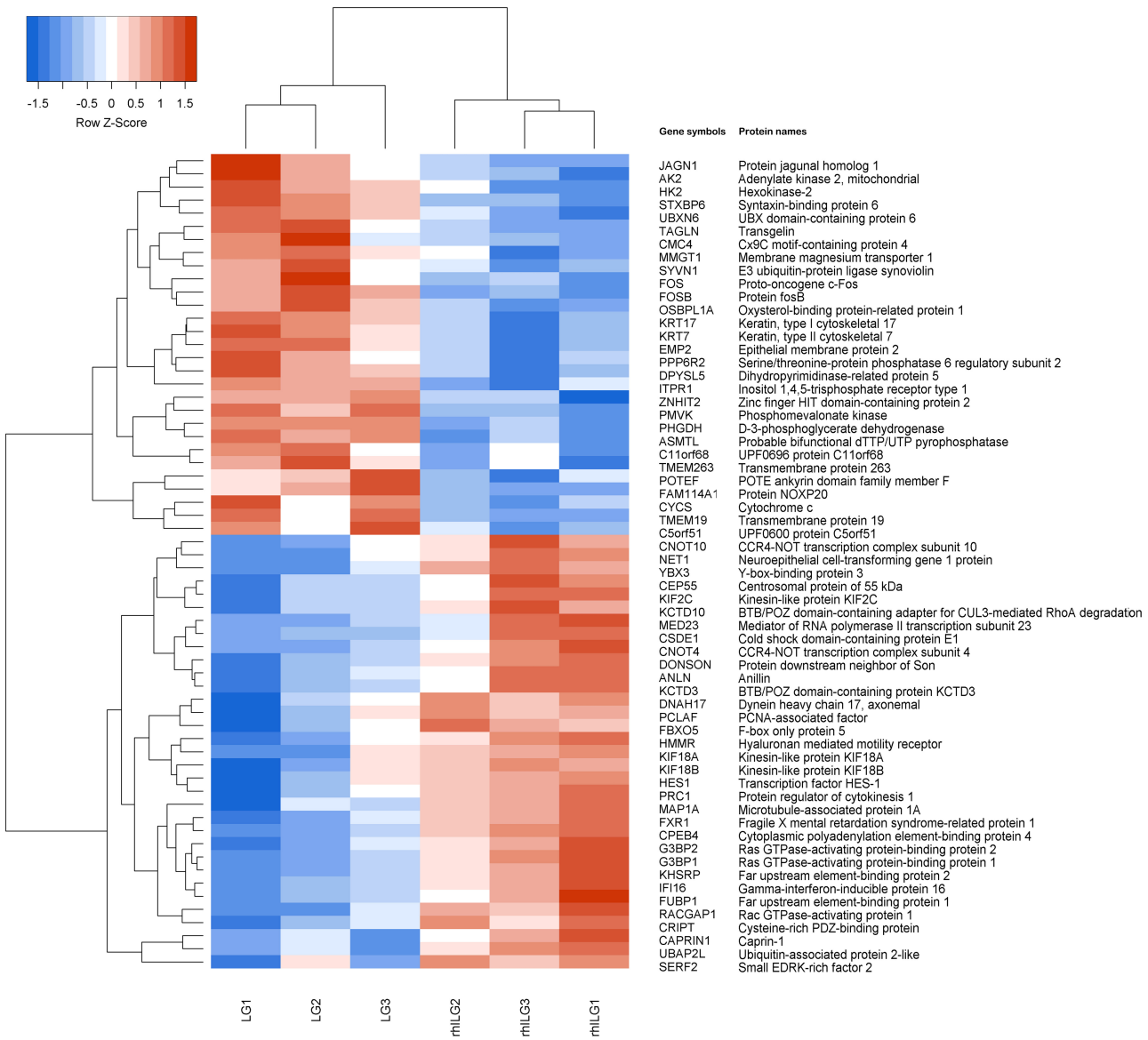


FIGURE 2 | Hierarchical clustering of the differentially expressed proteins in SaOS-2 osteosarcoma cells between the LG and rhILG. Each group contains three samples. The heatmap is a visualized demonstration of the protein distribution in different samples. The red color represents upregulation, and the blue color represents downregulation. The upper dendrogram illustrates the clustering analysis of different samples in different groups, and the left dendrogram shows the clustering analysis of different proteins in different samples. A total of 62 proteins were significantly changed between the LG and rhILG, including 33 upregulated proteins and 29 downregulated proteins ($p < 0.05$).

and $p = 0.0023$, respectively). No significant difference was found between the CG and rhIL-6G ($p = 0.087$). Tumor cell apoptosis in the different groups was then detected by flow cytometry. The average proportions of apoptotic cells were 7.13% ± 0.0041, 9.10% ± 0.0049, 70.4% ± 0.0037, and 60.5% ± 0.016 in the CG, rhIL-6G, LG, and rhILG for SaOS-2 cells and 9.40% ± 0.027, 11.37% ± 0.0079, 87.03% ± 0.0056, and 77.77% ± 0.0009 in the CG, rhIL-6G, LG, and rhILG for SOSP-9607 cells, respectively, which showed that rhIL-6 treatment significantly reduced the apoptosis of SaOS-2 and SOSP-9607 cells induced by the antineoplastic agent lobaplatin ($p < 0.05$) and enhanced the resistance of tumor cells to the drug. Hoechst 33342 was used to stain osteosarcoma cell nuclei of different groups, and the cells were observed by fluorescence microscopy. Cells in the LG exhibited obvious apoptotic morphology, which showed nuclear pyknosis and asymmetric chromatin condensation compared with cells in the rhILG, CG, and rhIL-6G (**Figure 1**).

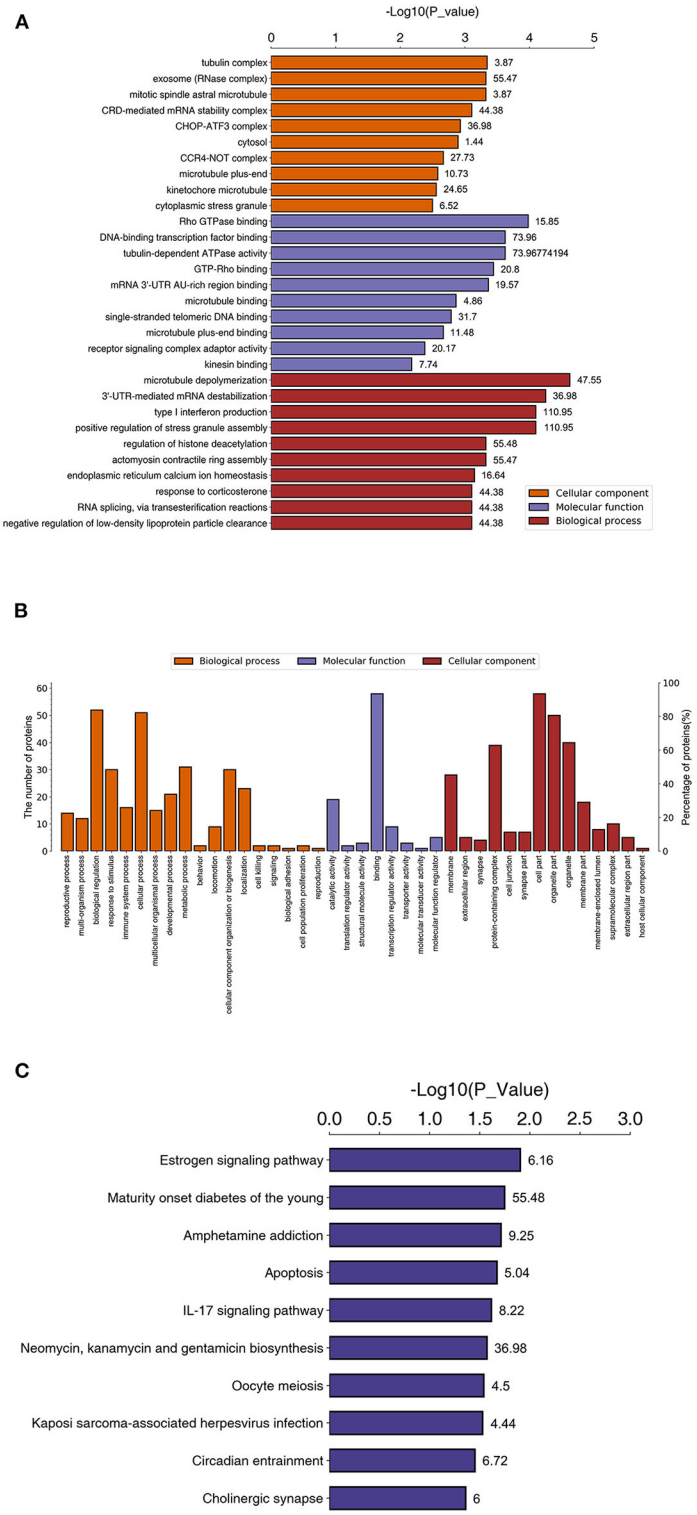


FIGURE 3 | Enriched GO and KEGG pathways between proteins in the LG and rhILG. Three replicates were performed for each group. **(A)** Top 10 enriched GO terms using Fisher's exact test for the biological process (BP), molecular function (MF), and cellular component (CC) categories. The vertical axis represents GO terms in each category, and the numbers beside the bars are enrichment factors, which represent the significance and reliability of proteins enriched in this item. The reliability of the proteins in an item was enhanced when the value increased. The horizontal axis demonstrates the –log 10 (*p*-value) of each item. **(B)** Total level 2 GO enrichment in the BP, MF, and CC categories. The vertical axes represent numbers of differentially expressed proteins that belong to a specific GO item (left) and the ratio of proteins in the item to the total differentially expressed proteins. **(C)** Significantly enriched pathways between the LG and rhILG using Fisher's exact test. The horizontal axis shows the significance of each pathway in the form of –log 10 (*p*-value). Numbers beside the bars are enrichment factors of each enriched pathway.

TABLE 1 | Enrichment of proteins and signaling pathways between the LG and rhILG groups based on GO and KEGG analysis.

Terms	Protein number	*p*-Value	FDR	Enrichment factor	Protein names (gene names)
GO					
(MF) Rho GTPase binding	3	3.592E-04	0.5243	15.85	Syntaxin-binding protein 6 (STXBP6); Anillin (ANLN); BTB/POZ domain-containing adapter for CUL3-mediated RhoA degradation protein 3 (KCTD10)
(CC) Tubulin complex	9	4.535E-04	0.1962	3.87	Dihydropyrimidinase-related protein 5 (DPYSL5); Cysteine-rich PDZ-binding protein (CRIPT); Rac GTPase-activating protein 1 (RACGAP1); Fragile X mental retardation syndrome-related protein 1 (FXR1); Protein regulator of cytokinesis 1 (PRC1); Kinesin-like protein (KIF18A); Kinesin-like protein (KIF2C); Kinesin-like protein (KIF18B); Microtubule-associated protein 1A (MAP1A)
(BP) Microtubule depolymerization	3	2.379E-05	1	47.55	Kinesin-like protein (KIF18A); Kinesin-like protein (KIF2C); Kinesin-like protein (KIF18B)
KEGG PATHWAYS					
Estrogen signaling pathway	3	0.01248	0.4391	6.16	Proto-oncogene (FOS); Keratin; type I cytoskeletal 17 (KRT17); Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)
Apoptosis	3	0.02134	0.4391	5.04	Proto-oncogene (FOS); Cytochrome c (CYCS); Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)
Amphetamine addiction	2	0.01942	0.4391	9.24	Protein fosB (FOSB); Proto-oncogene (FOS)
IL-17 signaling pathway	2	0.02428	0.4926	8.22	Protein fosB (FOSB); Proto-oncogene (FOS)
Oocyte meiosis	3	0.02874	0.6082	4.50	Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1); Cytoplasmic polyadenylation element-binding protein 4 (CPEB4); F-box only protein 5 (FBXO5)
Kaposi sarcoma-associated herpesvirus infection	3	0.02975	0.4961	4.44	Proto-oncogene (FOS); Cytochrome c (CYCS); Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)
Circadian entrainment	2	0.03528	0.4391	6.72	Proto-oncogene (FOS); Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)
Cholinergic synapse	2	0.04350	0.4391	6.00	Proto-oncogene (FOS); Inositol 1,4,5-trisphosphate receptor type 1 (ITPR1)

GO, gene ontology; KEGG pathways, Kyoto encyclopedia of genes and genomes pathways; MF, molecular functions; CC, cellular components; BP, biological processes; FDR, false discovery rate. GO and KEGG pathway enrichment were analyzed by the Fisher' exact test, and $p < 0.05$ were considered significant.

Hierarchical Cluster

Tandem mass tag liquid chromatography-tandem mass spectrometry was performed to identify some prominent proteins during rhIL-6 intervention before lobaplatin treatment in SaOS-2 osteosarcoma cells. A total of 1,313 differentially expressed proteins were identified and quantified, of which 62 proteins (33 upregulated; 29 downregulated) were found between the rhILG and LG (**Figure 2**), and 1,251 proteins (819 upregulated; 432 downregulated) were found between the LG and CG (**Supplementary Figure 1**). All the above proteins were either significantly up- or down-regulated by greater than a 1.2-fold change ($p < 0.05$). The heatmap of the hierarchical clustering demonstrated the differentially expressed proteins, which gives us a better visualization of the overall protein changes.

GO and KEGG Pathway Enrichment Analyses

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed between the rhILG and LG. The most enriched GO terms were annotated as Rho GTPase binding in the molecular function category (GO: 0017049, three proteins, enrichment factor = 15.85, p = 3.59165E-04), tubulin complex in the cellular compartment category (GO: 0045298, nine proteins, enrichment factor = 3.87, p = 4.53549E-04), and microtubule depolymerization in regard to the biological process category (GO: 0007019, three proteins,

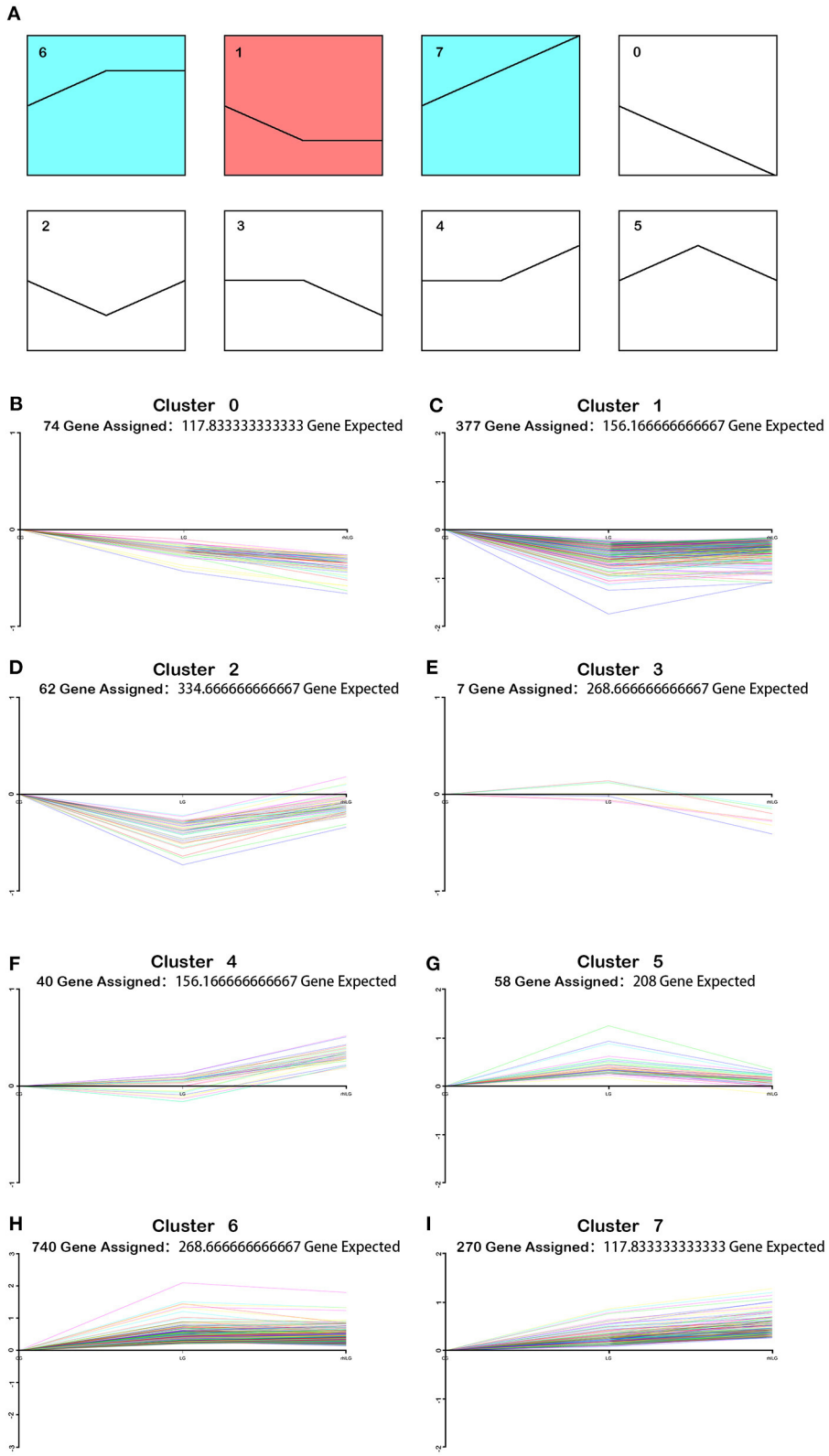


FIGURE 4 | Varied trends of differentially expressed proteins between the rhILG and LG using Short Time-series Expression Miner (STEM). **(A)** Proteins were grouped into eight clusters from 0 to 7. Clusters of similar colors represent similar trends. **(B–I)** Specific trends of different clusters; each line represents a different protein.

(Continued)

FIGURE 4 | (B) Seventy-four proteins were contained in cluster 0, in which proteins decreased in the LG compared with in the CG and decreased in the rhILG compared with the LG. **(C)** A total of 377 proteins were contained in cluster 1, in which proteins decreased in the LG compared with in the CG and slightly increased in the rhILG compared with in the LG. **(D)** Sixty-two proteins were contained in cluster 2, in which proteins decreased dramatically in the LG compared with in the CG and increased sharply in the rhILG compared with in the LG. **(E)** Seven proteins are contained in cluster 3, in which proteins increase or decrease slightly in the LG compared with in the CG and decrease sharply in the rhILG compared with in the LG. **(F)** Forty proteins were contained in cluster 4, in which proteins increase or decrease slightly in the LG compared with the CG and increase sharply in the rhILG compared with in the LG. **(G)** Fifty-eight proteins were contained in cluster 5, in which proteins increased sharply in the LG compared with in the CG and decreased dramatically in the rhILG compared with in the LG. **(H)** A total of 740 proteins were contained in cluster 6, in which proteins increased sharply in the LG compared with in the CG and showed no obvious changes in the rhILG compared with in the LG. **(I)** A total of 270 proteins were contained in cluster 7, in which proteins increased in the LG compared with in the CG and increased in the rhILG compared with in the LG.

enrichment factor = 47.55, p = 2.37872E-05) (**Figures 3A,B**, **Table 1**).

Kyoto Encyclopedia of Genes and Genomes pathway analysis demonstrated that the estrogen signaling pathway (three proteins, enrichment factor = 6.16, p = 0.01248), apoptosis pathway (three proteins, enrichment factor = 5.04, p = 0.02134), amphetamine addiction (two proteins, enrichment factor = 9.24, p = 0.01942), IL-17 signaling pathway (two proteins, enrichment factor = 8.22, p = 0.02428), oocyte meiosis (three proteins, enrichment factor = 4.50, p = 0.02874), Kaposi's sarcoma-associated herpesvirus infection (three proteins, enrichment factor = 4.44, p = 0.02975), circadian entrainment (two proteins, enrichment factor = 6.72, p = 0.03528), and cholinergic synapse (two proteins, enrichment factor = 6.00, p = 0.04350) were significantly enriched (**Figure 3C**, **Table 1**).

Cluster Analysis

To understand the role of rhIL-6 pretreatment on lobaplatin treatment of SaOS-2 osteosarcoma cells, the changing trends of differentially expressed proteins between the rhILG and LG were analyzed using Short Time-series Expression Miner (STEM) and divided into eight groups. The results showed that proteins in cluster 2 decreased sharply in the LG and increased in the rhILG, while proteins in cluster 5 increased in the LG and decreased in the rhILG. Proteins in these two groups showed that rhIL-6 exhibited the opposite effect compared to lobaplatin on SaOS-2 osteosarcoma cells (**Figure 4**). The proteins in the above two groups with a significance of lower than 0.05 are listed in **Supplementary Table 1**.

PPI Analysis

Protein–protein interaction (PPI) network analysis was performed to demonstrate the relationships of the differentially expressed proteins in cluster 2 and cluster 5 using both IntAct (http://www.ebi.ac.uk/intact/main.xhtml) and STRING (http://string-db.org/cgi/input.pl). As illustrated in IntAct, the most enriched highly connective proteins between the rhILG and LG were as follows: G3BP1 (UniProtKB-Q13283), proto-oncogene (UniProtKB-P01100), and gamma-interferon-inducible protein 16 (UniProtKB-Q16666), which have interactions with three significantly differentially expressed proteins and hFXR1p (UniProtKB-P51114), which interacts with two significantly differentially expressed proteins (**Figure 5A**, **Table 2**). Meanwhile, STRING suggested that the most enriched highly connective proteins between the rhILG and LG were as follows: G3BP1 (UniProtKB-Q13283), rac GTPase-activating protein 1 (UniProtKB-Q9H0H5), and protein regulator of cytokinesis 1 (UniProtKB-O43663), which each interact with five proteins; kinesin-like protein KIF2C (UniProtKB-Q99661) and ras GTPase-activating protein-binding protein 2 (UniProtKB-Q9UN86), which each interact with four proteins; and hFXR1p (UniProtKB-P51114), kinesin-like protein KIF18B (UniProtKB-Q86Y91), caprin-1 (UniProtKB-Q14444), and centrosomal protein of 55 kDa (UniProtKB-Q53EZ4), which interact with three significantly differentially expressed proteins in the rhILG and LG (**Figure 5B**, **Table 2**). Interestingly, we found that G3BP1 and hFXR1p exhibited a high connective degree in the above different enrichment methods.

PRM Validation of TMT-Based Results

Thirty-one differentially expressed proteins between the rhILG and LG, which were analyzed by GO and KEGG analyses and mainly in clusters 2 and 5, were selected for further verification using PRM quantitative analysis. The relative quantitative expression of each selected protein showed that 22 proteins exhibited similar trends as observed in the TMT results, while nine proteins did not demonstrate significant changes. As shown in **Figure 6**, the fold changes of G3BP1, FUBP1, and hFXR1p were verified to be significant in both the TMT and PRM analyses (**Figure 6**). However, the changes in nine proteins, including cold shock domain-containing protein E1, Rac GTPase-activating protein 1, and Kinesin-like protein, did not show statistical significance, as demonstrated by the TMT results (**Supplementary Table 2**).

Immunohistochemical and Immunofluorescence Staining of Clinical Specimens

Thirty osteosarcoma specimens, including 15 platinum-based chemotherapy-sensitive and 15 chemotherapy-resistant patients, were detected to assure the PRM results. The clinical data of the specimens are recorded and listed in **Supplementary Table 3**. The expression of G3BP1, FUBP1, and hFXR1p was strongly positive in 10, 9, and 10 platinum-based chemotherapy-resistant specimens, while in platinum-based chemotherapy-sensitive specimens, these proteins were negative, weakly positive, and at most moderately positive, respectively. The immunofluorescence staining results were in accordance with the immunohistochemistry staining results (**Figures 7A–D**).

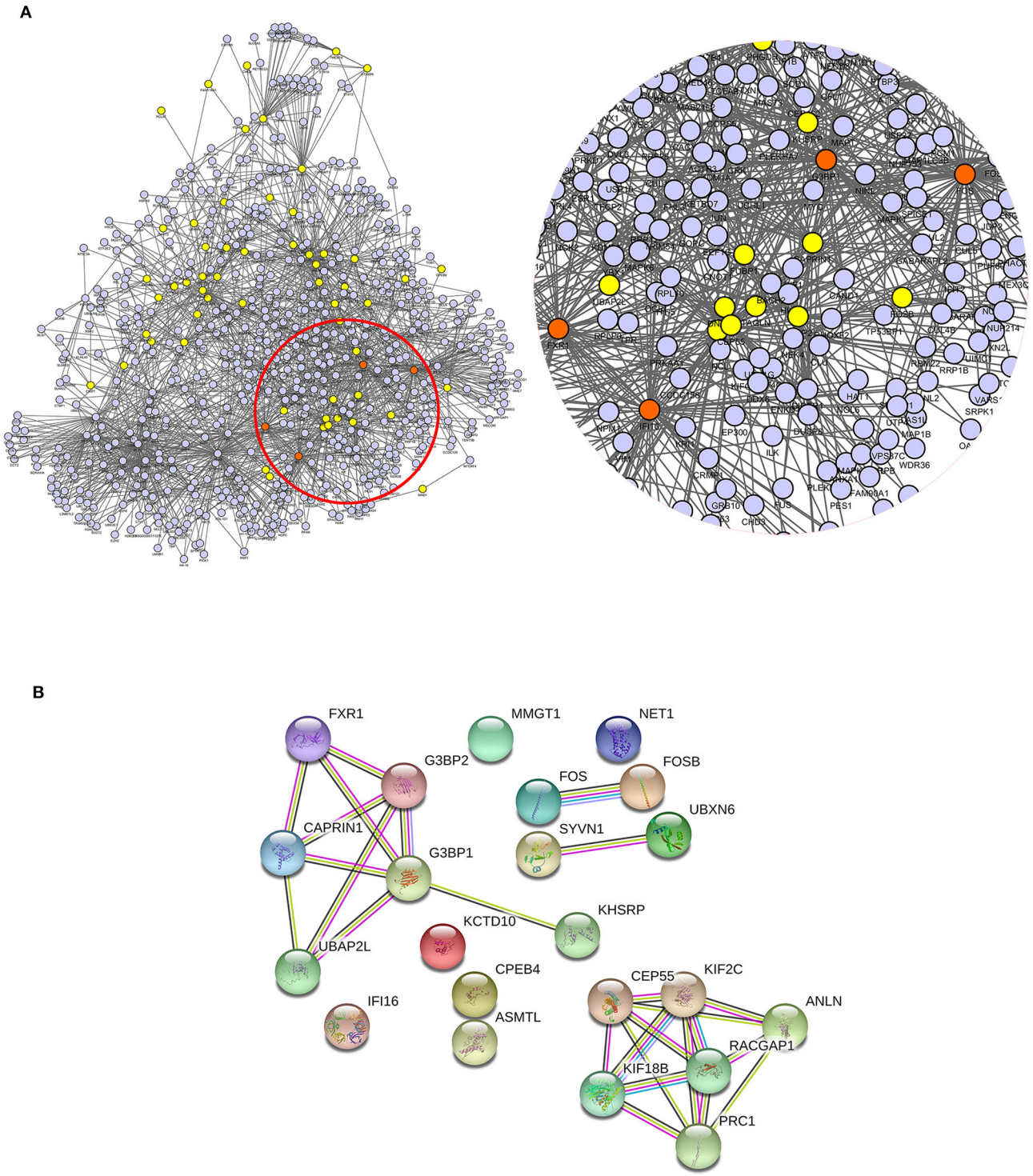


FIGURE 5 | Relationship of differentially expressed proteins in clusters 2 and 5 between the LG and rhILG. **(A)** PPI network analysis was performed using the IntAct database, which shows the direct interactions among the proteins as well as the relationship between these proteins and the reported linkers. The four most enriched node proteins with a high connecting degree are shown in orange. **(B)** The PPI network was constructed using the STRING database. Blue lines represent interactions from the curated database; purple lines represent interactions that were experimentally determined; dark green lines illustrate the gene neighborhood; light green lines represent text mining interactions; red lines exhibit gene fusions; and light violet lines represent protein homology. The two databases both show G3BP1 and hFXR1p as high connectivity proteins.

TABLE 2 | Comparison of nodes and highly connected proteins in the PPI in two databases.

Node proteins	Uniprot ID	Protein name	Gene name
IntAct			
Ras GTPase-activating protein-binding protein 1 (three connected)	Q16666	Gamma-interferon-inducible protein 16	IFI16
	Q14444	Caprin-1	CAPRIN1
	Q9UN86	Ras GTPase-activating protein-binding protein 2	G3BP2
Gamma-interferon-inducible protein 16 (three connected)	Q13283	Ras GTPase-activating protein-binding protein 1	G3BP1
	Q14444	Caprin-1	CAPRIN1
	Q16666	Gamma-interferon-inducible protein 16	IFI16
Proto-oncogene c-fos (three connected)	Q92945	Far upstream element-binding protein 2	KHSRP
	P53539	Protein fosB	FOSB
	P01100	Proto-oncogene c-Fos	FOS
Fragile X mental retardation syndrome-related protein 1 (two connected)	Q14157	Ubiquitin-associated protein 2-like	UBAP2L
	P51114	Fragile X mental retardation syndroame-related protein 1	FXR1
STRING			
Ras GTPase-activating protein-binding protein 1 (five connected)	P51114	Fragile X mental retardation syndrome-related protein 1	FXR1
	Q14444	Caprin-1	CAPRIN1
	Q9UN86	Ras GTPase-activating protein-binding protein 2	G3BP2
	Q14157	Ubiquitin-associated protein 2-like	UBAP2L
	Q92945	Far upstream element-binding protein 2	KHSRP
Rac GTPase-activating protein 1 (five connected)	Q53EZ4	Centrosomal protein of 55 kDa	CEP55
	Q99661	Kinesin-like protein KIF2C	KIF2C
	Q86Y91	Kinesin-like protein KIF18B	KIF18B
	Q9NQW6	Anillin	ANLN
	O43663	Protein regulator of cytokinesis 1	PRC1
Protein regulator of cytokinesis 1 (five connected)	Q9NQW6	Anillin	ANLN
	Q9H0H5	Rac GTPase-activating protein 1	RACGAP1
	Q99661	Kinesin-like protein KIF2C	KIF2C
	Q86Y91	Kinesin-like protein KIF18B	KIF18B
	Q53EZ4	Centrosomal protein of 55 kDa	CEP55
Kinesin-like protein KIF2C (four connected)	Q53EZ4	Centrosomal protein of 55 kDa	CEP55
	Q9NQW6	Anillin	ANLN
	Q86Y91	Kinesin-like protein KIF18B	KIF18B
	Q9H0H5	Rac GTPase-activating protein 1	RACGAP1
Ras GTPase-activating protein-binding protein 2 (four connected)	P51114	Fragile X mental retardation syndrome-related protein 1	FXR1
	Q13283	Ras GTPase-activating protein-binding protein 1	G3BP1
	Q14444	Caprin-1	CAPRIN1
	Q14157	Ubiquitin-associated protein 2-like	UBAP2L
Fragile X mental retardation syndrome—related protein 1 (three connected)	Q13283	Ras GTPase-activating protein-binding protein 1	G3BP1
	Q9UN86	Ras GTPase-activating protein-binding protein 2	G3BP2
	Q14444	Caprin-1	CAPRIN1
Kinesin-like protein KIF18B (three connected)	Q53EZ4	Centrosomal protein of 55 kDa	CEP55
	O43663	Protein regulator of cytokinesis 1	PRC1
	Q9H0H5	Rac GTPase-activating protein 1	RACGAP1
Centrosomal protein of 55 kDa (three connected)	Q99661	Kinesin-like protein KIF2C	KIF2C
	Q86Y91	Kinesin-like protein KIF18B	KIF18B
	Q9H0H5	Rac GTPase-activating protein 1	RACGAP1

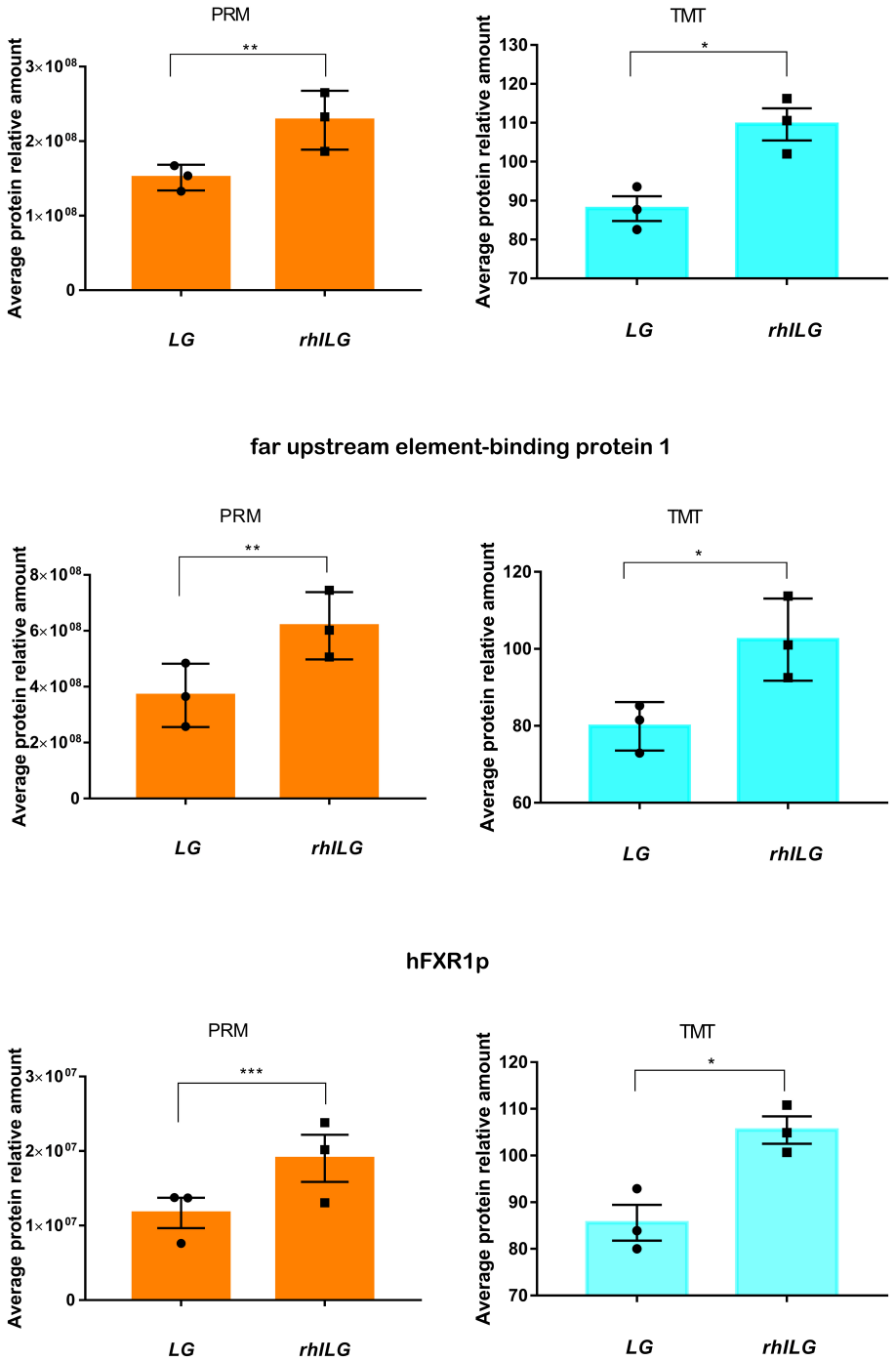


FIGURE 6 | Expression of three candidate proteins by TMT label-based LC-MS/MS and PRM validation. Each group contains three samples. Student's *t*-test. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

FUBP1 Knockdown Conferred Lobaplatin Sensitivity of Osteosarcoma SaOS-2 Cells

The three proteins' genes were separately silenced by siRNAs, and the viability of osteosarcoma SaOS-2 cells that were then treated with lobaplatin as described before was measured by CCK8 assay. The results showed that the IC50 values for lobaplatin were strikingly decreased in the FUBP1-silenced cells (RNAi#1, $p < 0.05$), while the IC50 values were not

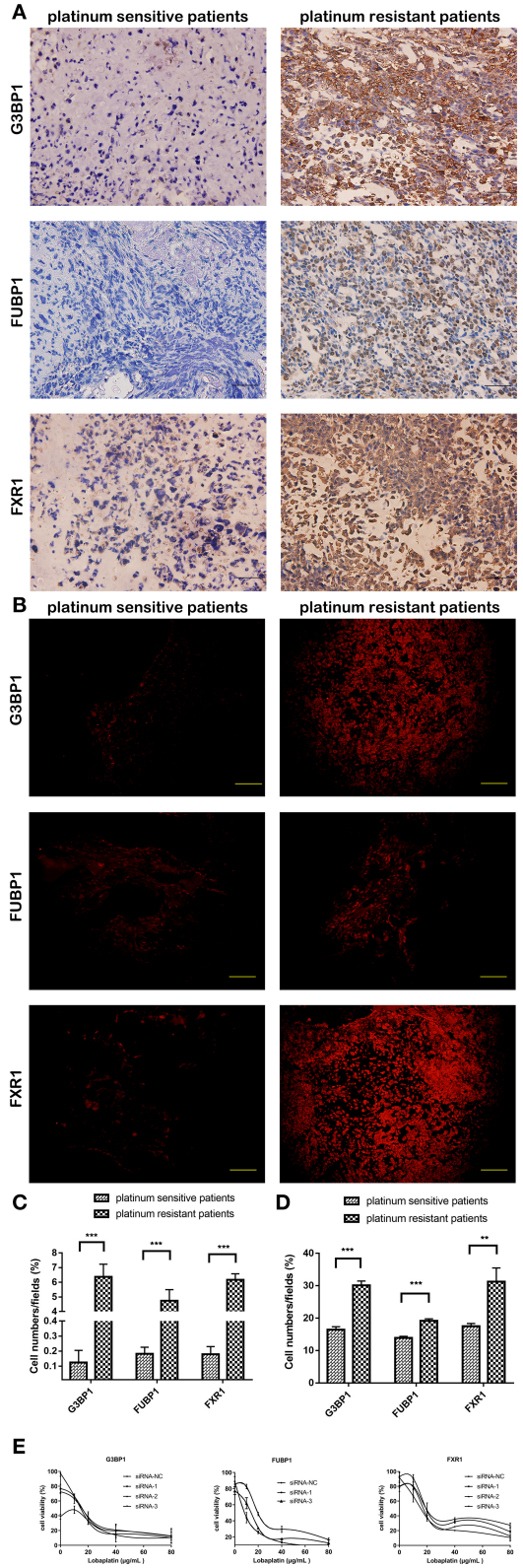


FIGURE 7 | Expression of three candidate proteins show difference between clinical platinum-sensitive and platinum-resistant specimens, while only

(Continued)

FIGURE 7 | knockdown of FUBP1 conferred lobaplatin sensitivity in osteosarcoma cells. **(A)** Representative immunohistochemistry staining of clinical specimens; scale bar = 50 μm (magnification, ×400). **(B)** Representative immunofluorescence staining of clinical specimens; scale bar = 100 μm (magnification, ×200). **(C)** Quantification of the expression of the three proteins in immunohistochemistry staining. **(D)** Quantification of the expression of the three proteins in immunofluorescence staining. **(E)** IC50 of lobaplatin in osteosarcoma SaOS-2 cells which were transfected with siRNAs targeting G3BP1, FUBP, and FXR1. Each bar represents the mean ± SEM of three independent experiments. $^{**}P < 0.01$, $^{***}P < 0.001$.

significantly changed in the G3BP1 and FXR1-silenced cells (**Figure 7E**).

DISCUSSION

Previous studies have indicated that exogenous and autocrine IL-6 confers chemotherapeutic resistance against platinum-based drugs (6, 15, 16), which is similar to the inference from our CCK-8 and flow cytometry assay in SaOS-2 and SOSP-9607 osteosarcoma cells. However, no comprehensive mechanisms have been explored from the protein perspective. Undoubtedly, proteins are ultimate performers of physiological function as well as a direct reflection of life activity, which is dynamic and flexible. Therefore, investigations into proteins and PPIs are conducive to understanding the whole story. In this study, we demonstrated that IL-6 intervention could decrease the sensitivity of SaOS-2 and SOSP-9607 osteosarcoma cells to lobaplatin. Additionally, a large-scale proteomic analysis among tumor cells in the CG, LG, and rhILG was performed using TMT labeling LC-MS/MS, which could overcome the shortcomings of traditional mass MS after 2D-PAGE and obtain more macromolecules and proteins.

This is the first study using TMT label-based LC-MS/MS technology on IL-6 intervention before osteosarcoma cells were treated with lobaplatin. In the present work, 1,251 proteins with significant differential expression were found between the LG and CG, and 62 proteins with significant differential expression were identified when comparing the LG and rhILG. Namely, there are a host of differentially expressed proteins when cells are treated with lobaplatin compared to control tumor cells, while no more than 100 proteins were differentially expressed between cells treated with lobaplatin and those pretreated with IL-6 before lobaplatin treatment. These results suggest that better homogeneity could be found in cell models than in tissues, which might show complicated heterogeneity; thus, we were able to acquire more accurate results instead of looking for a needle in a haystack.

GO and KEGG enrichment analyses suggested that IL-6 intervention before osteosarcoma cell treatment with lobaplatin significantly regulated Rho GTPase binding proteins and the tubulin complex as well as the estrogen signaling pathway, apoptosis pathway, and IL-17 signaling pathway, which is partly in accordance with our PPI analysis and the flow cytometry and Hoechst staining results. Regarding the estrogen signaling pathway, another study suggested that estrogen can restore tamoxifen sensitivity in breast cancer cells by inducing apoptosis

in tamoxifen-resistant cells (17), which explained the enrichment of the estrogen and apoptosis pathways by KEGG analysis. Hence, we speculate that the estrogen and apoptosis pathways may indicate potential therapeutic targets for lobaplatin-resistant osteosarcomas. IL-6 presumably decreased the chemosensitivity of osteosarcoma cells to lobaplatin through estrogen, apoptosis or IL-17 signaling, and proteins involved in the above pathways, such as proto-oncogene, inositol 1,4,5-trisphosphate receptor type 1, keratin, type I cytoskeletal 17, and cytochrome c, may represent other potential targets for lobaplatin-resistant osteosarcoma. This speculation needs to be verified by further experiments in a following study. Interestingly, we also found that two proteins are present in both the estrogen and apoptosis signaling pathways. This is the first report suggesting changes in estrogen as a possible pathway for IL-6-induced resistance to lobaplatin in an osteosarcoma cell model.

The candidate protein G3BP1 showed the highest connectivity with other differentially expressed proteins between the LG and rhILG using both IntAct and STRING. Ras GTPase-activating protein-binding protein 1 (G3BP1) is known to play an essential role in innate immunity as well as ras protein signal transduction. Moreover, it also promotes tumor progression and metastases (18–20). Additionally, it was reported that blocking the functions of ras GTPase-activating protein SH3 domain-binding protein (G3BP) can markedly suppress colon carcinoma HCT116 cell growth, and the downregulation of G3BP could enhance cisplatin-induced apoptosis (21). Although this protein was not enriched to the apoptosis pathway, it is related with cell growth and death to some extent. In our study, the expression of G3BP1 decreased in the LG compared with the CG and increased in rhIL-6-pretreated osteosarcoma cells compared with tumor cells treated with lobaplatin alone. The above finding suggests that changes in this protein are involved in the reduced susceptibility of osteosarcoma to lobaplatin.

Another protein, FUBP1, has been described as an oncoprotein in solid tumor entities as well as many other cancers, promoting tumorigenesis, proliferation, and metastasis of malignant cells (22–26). The silencing of FUBP1 could advance chemosensitivity to adriamycin in gastric cancer (27), and the expression of this protein in B-cell non-Hodgkin lymphoma is associated with cell adhesion-mediated drug resistance (28). Additionally, the UniProt database shows that it can transcriptionally activate the expression of cyclin-D2. In this study, PRM analysis verified that the expression of FUBP1 in rhILG cells was elevated compared with that in LG cells. Far upstream element-binding protein 1 expression was also related to the phenotype of reduced susceptibility to lobaplatin, indicating that it may contribute to reduced sensitivity induced by the cytokine IL-6.

Fragile X mental retardation syndrome-related protein 1 is a highly conserved RNA-binding protein known for its role in muscle development, inflammation, tumorigenesis, and metastatic behavior (29–32) that was enriched for the tubulin complex in the cellular compartment category (GO: 0045298) in our study. A genome-wide RNAi screen identified a host of potential drug resistance genes, including FXR1, by Attila (33). Moreover, Vaquero found that activation of FXR1 enhances chemoresistance of liver tumors against genotoxic compounds (34). Nevertheless, the role of hFXR1p in the drug resistance of osteosarcoma cells to lobaplatin, especially in the presence of IL-6, is unclear. Our proteomics-based study as well as PRM validation demonstrated that hFXR1p is a prominently increased protein in the rhILG compared with the LG, indicating that hFXR1p may be involved in lobaplatin resistance induced by exogenous IL-6 in osteosarcoma cells. The identification of the novel proteins above could lead to the discovery of new drug targets.

Immunohistochemical and immunofluorescence staining were further performed to verify the above three proteins in osteosarcoma patients who underwent platinum-based chemotherapy. Nearly 2/3 specimens from the chemotherapy-resistant group showed extremely strong positive expression of the three proteins, while negative to moderately positive expression was exhibited in the chemotherapy-sensitive group. However, the differential expression of the three proteins in clinical specimens only suggests a correlation between the proteins and the clinical phenomenon. Actually, they may contribute to drug resistance or they are the results of drug resistance.

siRNAs were synthesized to separately silence the three genes, *G3BP1*, *FUBP1,* and *FXR1*, and the sensitivity of osteosarcoma cells to lobaplatin were detected. Our study demonstrated that only FUBP1-silenced cells showed a decreased IC50 for lobaplatin. Combined with the proteomics and bioinformatics results, as well as the immunohistochemistry staining in clinical specimens, we can conclude that FUBP1 is a special protein that confers lobaplatin resistance in osteosarcoma cells. These findings collectively provide further proof of the role of FUBP1 in the drug susceptibility of osteosarcoma as well as the potential therapeutic value for increasing the sensitivity to lobaplatin in patients with osteosarcoma.

In conclusion, we noted that exogenous IL-6 intervention before lobaplatin treatment resulted in reduced sensitivity of osteosarcoma cells. TMT labeling LC-MS/MS and the subsequent PRM analysis reported herein point to three proteins, including G3BP1, FUBP1, and hFXR1p, as predictive markers to predict lobaplatin responsiveness and clinical outcomes. Although the immunohistochemistry verification was limited to a relatively smaller cohort due to the rarity of osteosarcoma patients who could meet the inclusion criteria, the above proteins present differential expression in platinum-resistant and platinum-sensitive specimens. Separate knockdown of the three genes verified that FUBP1 holds exciting potential as a new target protein to sensitize osteosarcoma cells to lobaplatin treatment. Further studies including larger patient cohorts are needed to clinically evaluate this protein.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (35) partner repository with the dataset identifier ProteomeXchange: PXD021564.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Tangdu Hospital, Fourth Military Medical University (permission number: TDLL-201903-30). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

QM conceived and designed the study. HW and BL analyzed the data and prepared the manuscript. YWu performed *in vitro* experiments in cell lines. KY collected clinical specimens and information. YWe and YL participated in IHC and IF staining. QM and PF provided financial support. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by National Natural Science Foundation of China (No. 81372297 and No. 81702660).

ACKNOWLEDGMENTS

We thank Yingxue Wang for her technical assistance in the experiments and her contribution to amending the language of the manuscripts.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.602712/full#supplementary-material

REFERENCES

1. Gianferante DM, Mirabello L, Savage SA. Germline and somatic genetics of osteosarcoma – connecting aetiology, biology and therapy. *Nat Rev Endocrinol.* (2017) 13:480–91. doi: 10.1038/nrendo.2017.16
2. Whelan JS, Davis LE. Osteosarcoma, chondrosarcoma, and chordoma. *J Clin Oncol.* (2018) 36:188–93. doi: 10.1200/JCO.2017.75.1743
3. Harrison DJ, Schwartz CL. Osteogenic sarcoma: systemic chemotherapy options for localized disease. *Curr Treat Options Oncol.* (2017) 18:24. doi: 10.1007/s11864-017-0464-2
4. Zhang Z, Wang X, Luo C, Zhu C, Wang K, Zhang C, et al. Dinuclear platinum(II) complexes with bone-targeting groups as potential anti-osteosarcoma agents. *Chem Asian J.* (2017) 12:1659–67. doi: 10.1002/asia.201700577
5. Oude Munnink T, van der Meer A, de Haan J, Touw D, van Kruchten M. Reversible impaired methotrexate clearance after platinum-based chemotherapy for osteosarcoma. *Ther Drug Monit.* (2019) 41:693–5. doi: 10.1097/FTD.0000000000000662
6. Pu YS, Hour TC, Chuang SE, Cheng AL, Lai MK, Kuo ML. Interleukin-6 is responsible for drug resistance and anti-apoptotic effects in prostatic cancer cells. *Prostate.* (2004) 60:120–9. doi: 10.1002/pros.20057
7. Ma H, Yan D, Wang Y, Shi W, Liu T, Zhao C, et al. Bazedoxifene exhibits growth suppressive activity by targeting interleukin-6/glycoprotein 130/signal transducer and activator of transcription 3 signaling in hepatocellular carcinoma. *Cancer Sci.* (2019) 110:950–61. doi: 10.1111/cas.13940
8. Culig Z, Puhr M. Interleukin-6 and prostate cancer: current developments and unsolved questions. *Mol Cell Endocrinol.* (2018) 462(Pt A):25–30. doi: 10.1016/j.mce.2017.03.012
9. Ham IH, Oh HJ, Jin H, Bae CA, Jeon SM, Choi KS, et al. Targeting interleukin-6 as a strategy to overcome stroma-induced resistance to chemotherapy in gastric cancer. *Mol Cancer.* (2019) 18:68. doi: 10.1186/s12943-019-0972-8
10. Zhang K, Dong C, Chen M, Yang T, Wang X, Gao Y, et al. Extracellular vesicle-mediated delivery of miR-101 inhibits lung metastasis in osteosarcoma. *Theranostics.* (2020) 10:411–25. doi: 10.7150/thno.33482
11. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods.* (2009) 6:359–62. doi: 10.1038/nmeth.1322
12. Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics.* (2006) 7:191. doi: 10.1186/1471-2105-7-191
13. Rodel C, Martus P, Papadoupolos T, Fuzesi L, Klimpfinger M, Fietkau R, et al. Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer. *J Clin Oncol.* (2005) 23:8688–96. doi: 10.1200/JCO.2005.02.1329
14. Ma Q, Zhang Y, Liu T, Jiang K, Wen Y, Fan Q, et al. Hypoxia promotes chemotherapy resistance by down-regulating SKA1 gene expression in human osteosarcoma. *Cancer Biol Ther.* (2017) 18:177–85. doi: 10.1080/15384047.2017.1294285
15. Wang Y, Niu XL, Qu Y, Wu J, Zhu YQ, Sun WJ, et al. Autocrine production of interleukin-6 confers cisplatin and paclitaxel resistance in ovarian cancer cells. *Cancer Lett.* (2010) 295:110–23. doi: 10.1016/j.canlet.2010.02.019
16. Duan Z, Lamendola DE, Penson RT, Kronish KM, Seiden MV. Overexpression of IL-6 but not IL-8 increases paclitaxel resistance of U-2OS human osteosarcoma cells. *Cytokine.* (2002) 17:234–42. doi: 10.1006/cyto.2001.1008
17. Mansouri S, Farahmand L, Hosseinzade A, Eslami SZ, Majidzadeh AK. Estrogen can restore Tamoxifen sensitivity in breast cancer cells amidst the complex network of resistance. *Biomed Pharmacother.* (2017) 93:1320–5. doi: 10.1016/j.biopha.2017.07.057
18. Wang Y, Fu D, Chen Y, Su J, Wang Y, Li X, et al. G3BP1 promotes tumor progression and metastasis through IL-6/G3BP1/STAT3 signaling axis in renal cell carcinomas. *Cell Death Dis.* (2018) 9:501. doi: 10.1038/s41419-018-0504-2
19. Dou N, Chen J, Yu S, Gao Y, Li Y. G3BP1 contributes to tumor metastasis via upregulation of Slug expression in hepatocellular carcinoma. *Am J Cancer Res.* (2016) 6:2641–50.
20. Somasekharan SP, El-Naggar A, Leprivier G, Cheng H, Hajee S, Grunewald TG, et al. YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. *J Cell Biol.* (2015) 208:913–29. doi: 10.1083/jcb.201411047
21. Zhang H, Zhang S, He H, Zhao W, Chen J, Shao RG. GAP161 targets and downregulates G3BP to suppress cell growth and potentiate cisplaitin-mediated cytotoxicity to colon carcinoma HCT116 cells. *Cancer Sci.* (2012) 103:1848–56. doi: 10.1111/j.1349-7006.2012.02361.x
22. Jiang P, Huang M, Qi W, Wang F, Yang T, Gao T, et al. FUBP1 promotes neuroblastoma proliferation via enhancing glycolysis-a new possible marker of malignancy for neuroblastoma. *J Exp Clin Cancer Res.* (2019) 38:400. doi: 10.1186/s13046-019-1414-6
23. Debaize L, Troadec MB. The master regulator FUBP1: its emerging role in normal cell function and malignant development. *Cell Mol Life Sci.* (2019) 76:259–81. doi: 10.1007/s00018-018-2933-6
24. Debaize L, Jakobczyk H, Avner S, Gaudichon J, Rio AG, Serandour AA, et al. Interplay between transcription regulators RUNX1 and FUBP1 activates an enhancer of the oncogene c-KIT and amplifies cell proliferation. *Nucleic Acids Res.* (2018) 46:11214–28. doi: 10.1093/nar/gky756
25. Hoang VT, Verma D, Godavarthy PS, Llavona P, Steiner M, Gerlach K, et al. The transcriptional regulator FUBP1 influences disease outcome in murine and human myeloid leukemia. *Leukemia.* (2019) 33:1700–12. doi: 10.1038/s41375-018-0358-8
26. Zhong Q, Liu ZH, Lin ZR, Hu ZD, Yuan L, Liu YM, et al. The RARS-MAD1L1 fusion gene induces cancer stem cell-like properties and therapeutic resistance in nasopharyngeal carcinoma. *Clin Cancer Res.* (2018) 24:659–73. doi: 10.1158/1078-0432.CCR-17-0352

27. Zhao D, Zhang Y, Song L. MiR-16-1 Targeted silences far upstream element binding protein 1 to advance the chemosensitivity to adriamycin in gastric cancer. *Pathol Oncol Res.* (2018) 24:483–8. doi: 10.1007/s12253-017-0263-x
28. Huang Y, Xu X, Ji L, Wang Y, Wang S, Tang J, et al. Expression of far upstream element binding protein 1 in Bcell nonHodgkin lymphoma is correlated with tumor growth and celladhesion mediated drug resistance. *Mol Med Rep.* (2016) 14:3759–68. doi: 10.3892/mmr.2016.5718
29. McClure JJ, Palanisamy V. Muscle-Specific FXR1 isoforms in squamous cell cancer. *Trends Cancer.* (2019) 5:82–4. doi: 10.1016/j.trecan.2018.12.001
30. Herman AB, Vrakas CN, Ray M, Kelemen SE, Sweredoski MJ, Moradian A, et al. FXR1 is an IL-19-responsive RNA-binding protein that destabilizes pro-inflammatory transcripts in vascular smooth muscle cells. *Cell Rep.* (2018) 24:1176–89. doi: 10.1016/j.celrep.2018.07.002
31. Fan Y, Yue J, Xiao M, Han-Zhang H, Wang YV, Ma C, et al. FXR1 regulates transcription and is required for growth of human cancer cells with TP53/FXR2 homozygous deletion. *Elife.* (2017) 6:e26129. doi: 10.7554/eLife.26129
32. Qie S, Majumder M, Mackiewicz K, Howley BV, Peterson YK, Howe PH, et al. Fbxo4-mediated degradation of Fxr1 suppresses tumorigenesis in head and neck squamous cell carcinoma. *Nat Commun.* (2017) 8:1534. doi: 10.1038/s41467-017-01199-8
33. Seyhan AA, Varadarajan U, Choe S, Liu W, Ryan TE. A genome-wide RNAi screen identifies novel targets of neratinib resistance leading to identification of potential drug resistant genetic markers. *Mol Biosyst.* (2012) 8:1553–70. doi: 10.1039/c2mb05512k
34. Vaquero J, Briz O, Herraez E, Muntane J, Marin JJ. Activation of the nuclear receptor FXR enhances hepatocyte chemoprotection and liver tumor chemoresistance against genotoxic compounds. *Biochim Biophys Acta.* (2013) 1833:2212–9. doi: 10.1016/j.bbamcr.2013.05.006
35. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* (2019) 47:D442–50. doi: 10.1093/nar/gky1106

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 Wang, Li, Yan, Wu, Wen, Liu, Fan and Ma. 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: 01 April 2021
doi: 10.3389/fcell.2021.568660

Transcriptional Repression of MFG-E8 Causes Disturbance in the Homeostasis of Cell Cycle Through DOCK/ZP4/STAT Signaling in Buffalo Mammary Epithelial Cells

Arvind K. Verma†, Syed A. Ali†‡, Parul Singh, Sudarshan Kumar and Ashok K. Mohanty‡*

Proteomics and Cell Biology Lab, Animal Biotechnology Center, ICAR-National Dairy Research Institute, Karnal, India

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Zhi Chen, Yangzhou University, China
Elrashdy M. Redwan, King Abdulaziz University, Saudi Arabia

***Correspondence:**
Ashok K. Mohanty
ashokmohanty1@gmail.com

†These authors have contributed equally to this work

‡ORCID:
Ashok K. Mohanty
orcid.org/0000-0002-9420-8966
Syed A. Ali
orcid.org/0000-0003-3024-9379

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

Received: *01 June 2020*
Accepted: *08 March 2021*
Published: *01 April 2021*

Citation:
Verma AK, Ali SA, Singh P, Kumar S and Mohanty AK (2021) Transcriptional Repression of MFG-E8 Causes Disturbance in the Homeostasis of Cell Cycle Through DOCK/ZP4/STAT Signaling in Buffalo Mammary Epithelial Cells. Front. Cell Dev. Biol. 9:568660. doi: 10.3389/fcell.2021.568660

The mammary gland is a unique apocrine gland made up of a branching network of ducts that end in alveoli. It is an ideal system to study the molecular mechanisms associated with cell proliferation, differentiation, and oncogenesis. MFG-E8, also known as Lactadherin, is a vital glycoprotein related to the milk fat globule membrane and initially identified to get secreted in bovine milk. Our previous report suggests that a high level of MFG-E8 is indicative of high milk yield in dairy animals. Here, we showed that MFG-E8 controls the cell growth and morphology of epithelial cells through a network of regulatory transcription factors. To understand the comprehensive action, we downregulated its expression in MECs by MFG-E8 specific shRNA. We generated a knockdown proteome profile of differentially expressed proteins through a quantitative iTRAQ experiment on a high-resolution mass spectrometer (Q-TOF). The downregulation of MFG-E8 resulted in reduced phagocytosis and cell migration ability, whereas it also leads to more lifespan to knockdown vis-a-vis healthy cells, which is confirmed through BrdU, MTT, and Caspase 3/7. The bioinformatics analysis revealed that MFG-E8 knockdown perturbs a large number of intracellular signaling, eventually leading to cessation in cell growth. Based on the directed network analysis, we found that MFG-E8 is activated by CX3CL1, TP63, and CSF2 and leads to the activation of SOCS3 and CCL2 for the regulation of cell proliferation. We further proved that the depletion of MFG-E8 resulted in activated cytoskeletal remodeling by MFG-E8 knockdown, which results in the activation of three independent pathways ZP4/JAK-STAT5, DOCK1/STAT3, and PIP3/AKT/mTOR. Overall, this study suggests that MFG-E8 expression in mammary epithelial cells is an indication of intracellular deterioration in cell health. To date, to the best of our knowledge, this is the first study that explores the downstream targets of MFG-E8 involved in the regulation of mammary epithelial cell health.

Keywords: milk fat globule epithelial growth factor-8, cell proliferation, quantitative proteomics, cytoscape, KEGG, oncogenesis

INTRODUCTION

Milk fat globule-EGF-factor is a 72-kDa secreted glycoprotein initially recognized in MFGM released into milk by MECs. It was identified to function as a bridging protein for apoptotic cells and phagocytes, leading to the subsequent engulfment and clearance of apoptotic cells. However, several other studies report that MFG-E8 is a multifunctional molecule shown to be released by a variety of cell types such as macrophages, immature dendrocytes, myoepithelial, endothelial, retinal, intestinal epithelial cells (Ceriani et al., 1983; Raymond et al., 2009; Zhou et al., 2018; Chopra et al., 2020). It was also found to be expressed at high levels in many tumor types (Carmon et al., 2002; Neutzner et al., 2007; Ko et al., 2020; Yamazaki et al., 2020). Hitherto, the most critical described function of MFG-E8 is to regulate immune homeostasis through the phagocytosis of apoptotic cells by signaling through $\alpha_v\beta_{3-5}$ integrins linking phosphatidylserine at the surface of membrane vesicles (Oshima et al., 2002) of apoptotic cells (Hanayama et al., 2002; Lotfan et al., 2018; Peterman et al., 2019).

During the inflammatory process, it mediates phagocytosis induced regulatory T cell response (Jinushi et al., 2007) and Mfge8$^{-/-}$ mice develop the spontaneous onset of lupus-like disease and glomerulonephritis (Hanayama et al., 2004). In addition to the immunological process, it also involves multiple regulatory functions, including in humans, SNP in MFG-E8 is associated with SLE (Hu et al., 2009). This protein shares structural domain homology with Del-1 (developmental endothelial locus 1), constituting a two-gene family of $\alpha_v\beta_3$ integrin ligands (Hidai et al., 1998). Leading to the binding of MFG-E8 to $\alpha_v\beta_3/\beta_5$ on vascular endothelial cells promotes VEGF-driven neovascularization (Silvestre et al., 2005) and knocks down suppresses glioma progression (Wu et al., 2020). Therefore, it drives us to study the impact of MFG-E8 for other functions, including the role in the crucial cellular proliferation and homeostasis through phagocytosis (Li et al., 2019).

Importantly, numerous investigations have demonstrated that MFG-E8 plays an essential role in mammary gland development, essentially harmonizing post-lactational mammary organ remodeling (Stubbs et al., 1990; Lönnerdal, 2003; Atabai et al., 2005; Hanayama and Nagata, 2005). The removal of apoptotic cells is necessary for developmental stages and cleansing the mammary gland during lactation from unwanted pathogens. In turn, these studies highlighted an essential function of MFG-E8 in maintaining homeostasis of the mammary gland stages for lactation and involution function (Stubbs et al., 1990; Lönnerdal, 2003; Atabai et al., 2005; Hanayama and Nagata, 2005). It was also reported that MFG-E8 transcripts increase in the mammary gland during pregnancy to lactation, suggesting its typical role during the glandular developmental transition (Oshima et al., 1999). Lately, we also identified in our DiGE based proteomics dataset that MFG-E8 is up-regulated in MECs isolated from the high milk yielding cows (Janjanam et al., 2014). Despite the several reports emphasizing the association between this molecule and various cellular physiologies, the exact mechanism, its intracellular targets, and downstream signaling circuits in MECs are not known. The interpretation of MFG-E8 mediated regulatory control on homeostasis in healthy and diseased conditions remains poorly understood.

Therefore, the present study aims to understand the essential molecular shifts under MFG-E8 signaling. In this work, we observed the effect of MFG-E8 repression on cellular physiology using iTRAQ based high throughput proteomics technique and respective biological assays. We determined the linkage of MFG-E8 to maintain cell cycle and cellular proliferation through the ZP4/JAK-STAT5, DOCK1/STAT3, and PIP3/AKT/mTOR pathways.

Abbreviations: Q-TOF, quadrupole time-of-flight; MFG-E8, milk fat globule-EGF factor 8 protein; MECs, mammary epithelial cells; MFGM, milk fat globules membrane; SNP, single nucleotide polymorphism; SLE, systemic lupus erythematosus; DiGE, difference gel electrophoresis; TFs, transcription factors; BoMac, the bovine macrophage; BuMEC, buffalo mammary epithelial cell; KD_MEC, KD BuMEC_MFG-E8 cell line; GFP, green fluorescent protein; DEPs, differentially expressed proteins; GEF, guanine nucleotide exchange factor.

MATERIALS AND METHODS

Cell Culture and Stable Knockdown of MFG-E8 Expression in BuMEC Cells

The BoMac cell line was the generous gift by Prof. J. R Stabel. The cells were cultured in RPMI medium with 10% (FBS) (Cat. No. A4766801, Gibco BRL, United States). The BuMEC was previously established in the lab [ExPASy accession: Buffalo 2012 (CVCL_M445)] and used for the experimental purpose. Briefly, cells were cultured in DMEM/F12 media (Cat. No. 11330032, Gibco BRL, United States) supplemented with 8% (FBS) (Cat. No. A4766801, Gibco BRL, United States), 5 μg/ml bovine insulin (Cat. No. I-035, Sigma, United States), 1 μg/ml hydrocortisone (Cat. No. H0888, Sigma, United States), 10 ng/ml EGF (Cat. No. E4127, Sigma, United States), 100 U/ml penicillin, and 5 μg/ml streptomycin (Cat. No. P4333, Sigma, United States) were incubated with 5% CO_2 at 37°C in a humidified environment. The confluent monolayer after 4–5 days were sub-cultured at split ratio of 1:3 by seeding in 1×10^6 cells/flasks (25 cm^2 flasks) by trypsinization (0.5% trypsin and 0.05% EDTA). The fresh medium was changed in ever two alternate days regularly.

The mfge8 gene shown in **Supplementary Figure 1** was used for the preparation of shRNA targets. The designed three shRNA MFG-E8_1 (TCCCACAAGAAGAACATATTT), MFG-E8_2 (CGGTCAGGAGATAAGATATTT) and MFG-E8_3 (CAACAGCGGCCTGAAGATTAA) oligonucleotides were annealed in pLKO.1-puro-CMV-tGFP Vector backbone individually and used for the transfection. For making the stable knockdown BuMECs (3×10^5) cell line, initially, cells were cultured in 24 well plate at 70% confluence, transfected with 3 μg of a non-target scramble and targeted shRNA pLKO.1-puro-CMV-tGFP Recombinant plasmid using lipofectamine 2000 (Cat. No. 11668027, Invitrogen Thermo Fisher Scientific, Waltham, MA, United States) using manufacturer protocol. After 24 h, cells were shifted to complete medium containing 50 μg/mL of puromycin (Cat. No. A1113803, Invitrogen Thermo Fisher Scientific, Waltham, MA, United States) as a selective agent. After 48 h of transfection, cells were trypsinized and

seeded at low density in puromycin-containing media and allowed to grow with regular media changed every 48 h. Colonies of transfected cells were selected and further propagated in the selection medium by retrograde passaging using clonal dilution till a stably transfected KD BuMEC_MFG-E8 cell line (KD_MEC) was established. The expression of GFP protein was observed under fluorescent microscope (NikonTE2000, Japan). The effect of the knockdown was confirmed by analyzing the low-level expression of MFG-E8 at different passages by quantitative real-time PCR. The cell line stably transfected with MFG-E8 shRNA pLKO.1-puro-CMV-tGFP was designated as KD_MEC_MFG-E8 (KD_MEC) and the control transfected with scramble pLKO.1-puro-CMV-tGFP plasmids were named as Puro_MEC_MFG-E8 (WT_MEC).

Sample Preparation for Proteomic Analysis

For the preparation of the proteomics sample, KD_MEC_MFG-E8 (KD_MEC) and Puro_MEC_MFG-E8 cells (WT_MEC) were cultured for 80% confluence, washed three times with PBS. The cells were harvested using 500 μL of lysis buffer [50 mM Tris-Cl (pH7.8), 0.3% SDS, 200 mM DTT, 1 mM PMSF, 1 mM EDTA] containing a protease inhibitor cocktail (Sigma-Aldrich) and vortexed. Subsequently, sonication was performed using the 10 cycles (pulse ON 5 s and OFF 7 s) with an interval of 120 s. The undisrupted cells were removed by centrifugation at 2500 × *g* for 15 min, and the supernatant was transferred to another tube. Proteins were precipitated by adding 10 volumes of cold acetone followed by overnight incubation at −20°C. The precipitated proteins were collected by centrifugation at 12,000 × *g* for 20 min. The protein concentration in the cell-free extract was determined using a 2D Quant kit (GE Healthcare, Life Sciences), according to manufacturer protocol.

Protein Digestion, iTRAQ Labeling, and RP-HPLC Fractionation

The iTRAQ labeling was performed as per the manufacturer protocol (AB Sciex, Marsh Rd., Foster City, CA, United States). Briefly, the lyophilized protein sample (100 μg) from each group was resuspended in 20 μl of 2% sodium dodecyl sulfate (SDS) in 1 M triethylammonium bicarbonate (TEAB). The samples then reduced in 50 mM tris-(2-carboxyethyl) phosphine (TCEP) for 1 h at 56°C and cysteine-blocked with 200 mM iodoacetamide (IAA) at room temperature in the dark for 30 min. The proteins were digested using trypsin (Mass Spec Grade, Promega, Madison, WI, United States) at 1:80 (trypsin: protein ratio) for 16 h at 37°C. The prepared peptides were iTRAQ labeled in duplicate pairs with KD_MEC_MFG-E8 (KD1-114, and KD2-116) and Puro_MEC_MFG-E8 (WT1-115, and WT2-117) for 1 h at room temperature. The reaction was stopped by adding deionized water, and the labeled samples were pooled. The final combined sample was fractionated by RP- HPLC (Agilent 1100) using an analytical column [Grace Smart RP, C18 (150 × 4.6 mm), 5 μm]. The mobile phase A (10 mM TEAB) and B (10 mM TEAB in 90% ACN) was used, respectively, for the peptides separation using following linear gradient: 0 to 2% in 5 min, 2 to 60% in 60 min, 60 to100% in 25 min and then 100 to 2% in 2 min. The collected 120 fractions were concatenated into 30 fractions, again lyophilized and reconstituted in 0.1% formic acid in water and subjected to desalting using C18 zip tip (Millipore, United States). Eluted peptides were lyophilized and re-dissolved in 0.4% FA and immediately used for MS/MS spectra generation.

Electrospray Ionization Tandem Mass Spectrometry LC-MS/MS Analysis

The reconstituted peptides were used for shotgun LFQ and iTRAQ proteomics experiments. The peptides were separated using nano-LC (Nano-Advance, Bruker, Germany) through a trap column (Bruker Magic C18 AQ, 0.1 × 20 mm, 3 μm, 200 Å) and a nano-analytical column (Bruker Magic C18 AQ, 0.1 × 150 mm, 3 μm, 200 Å) coupled with captive ion source (Bruker Captive Spray tip) spray in Maxis-HD qTOF (Bruker, Germany) mass spectrometer (MS) for identification with high mass accuracy and sensitivity. The elution was performed with a flow rate of 400 nL/min in a continuous gradient of 5–75% acetonitrile over 135 min. In the solvent system; Solvent A was 100% water in 0.1% formic acid, and solvent B was 100% acetonitrile in 0.1% formic acid. For acquisition, the data-dependent mode was used in mass spectrometer operated into automatically switch between MS and MS/MS acquisition. The precursor ion MS spectra scan range of 300–1400 (m/z) was used in the Q-TOF with resolution $R = 75,000$. The six most abundant precursor ions were searched for detection of different masses in the acquisition method and selected for fragmentation using collision-induced dissociation (CID) with a fixed cycle time of 3 s along with 2 min of release for exclusion filter (otof processing software, Bruker Daltonics). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1] partner repository with the dataset identifier PXD022445.

Data Processing for iTRAQ Analysis

With the above mentioned same parameter, the iTRAQ analysis was performed in the Maxquant with Andromeda environment. Simultaneously, the data were also analyzed using the TPP pipeline. For TPP analysis, the obtained otof generated raw (.d) files were converted to mzML format using MSconvertGUI using the default parameters. The converted mzML files were searched using the Trans-Proteomic Pipeline version 5.1.0 released on 2017-11-03 to in-house combined UniProt (Cow), and *Bubalus bubalis* (Buffalo) database together with common contaminant sequences were provided for MS/MS spectra search. The database was also appended with an equal number of decoy sequences (reversed proteins sequences from the original database). Initially, for the analysis, the peptide assignments were performed using multiple search engines using X! Tandem (with the *k*-score plug-in) (MacLean et al., 2006), SpectraST, and Comet. The multiple search engines parameters were adapted from our previously published work (Suhail et al., 2019). For all the search engines, the search parameters included semi-digested LysC and trypsin with two allowed miss cleavage, static modifications carbamidomethyl (Cys), and iTRAQ reagents (N-terminus and

Lys), and dynamic modification oxidation (Met), Gln-pyro Glu and Glu-pyro Glu remaining parameter were kept as default. The minimum peptide length parameter was set to seven amino acid residue. Further, PeptideProphet and ProteinProphet algorithms were used in the pipeline to compute the probabilities score for both individually searched peptides and the respective proteins. The accurate mass model in PeptideProphet was used for high confidence peptide identifications to boost the probability of peptides (Deutsch et al., 2010). Another protein validation step was executed using both Peptide Prophet and Protein Prophet scores, where the protein was authenticated if it contains minimum two top-ranked peptides with each peptide probability score above 95% (**Supplementary Figure 1**).

All the search engine results were merged and validated using another computational method, termed iProphet. This method takes as the input of PeptideProphet spectrum-level results from multiple LC-MS/MS runs and then computes a new probability at the level of a unique peptide sequence [or protein sequence (Nesvizhskii et al., 2003)]. This framework allows for the combination of results from multiple search tools. It takes into account other supporting factors, including the number of sibling experiments identifying the same peptide ions, the number of replicate ion identifications, sibling ions, and sibling modification states. A model of iProphet performance for the number of correct identifications *versus* error is shown in **Supplementary Figure 1**. An iProphet probability of more than 0.95 was used as the cutoff for the final identification of the protein. For protein quantitation, $\geq$2 unique peptides per protein were considered to ensure high-quality quantitation.

Bioinformatics Analysis and Network Construction

The protein–protein interaction networks were constructed utilizing DEPs, to examine the complex interaction and prediction of pathways. We retrieved the whole protein sequences using an in-house made python script from the UniProt database for all proteins, and these proteins were mapped for protein–protein interactions networks using STRING v 10.0 [highest confidence (0.90)], Kyoto Encyclopedia of Genes and Genomes (KEGG)[1], GeneMANIA, and Reactome. Further, Cytoscape 3.2.1 was used to visualize networks and were manually for biologically relevant processes. We used R package ggplot2 for creating the volcano plot while PCA plots were prepared using the factoextra package. Two separate PCA plots were created, first is the individual group component analysis for discrimination of Wild type cells and Knockdown cells (i.e., complete set of experiments). Second is the biplot of variables PCA, and it takes into account the consideration of all quantified proteins (i.e., complete set of proteins).

To further elucidate the biological inference of the enrichment terms and associated pathways, we employed Cytoscape plug-in ClueGO. It combines the pathways and gene ontology terms in functional networks, which represent the significance of the relationship between the pathways. The degree of nodes is designated by the number of edges it is connected for a network. It functions to illustrate the relationships between the Gene Ontology (GO) terms based on their overlapping genes. The term-term similarity matrix was calculated to establish the connection strength between the terms based on the highest significance. This study uses the following parameters kappa score of 0.4, GO tree interval 3 minimum level, 8 maximum levels, GO term/pathway selection four minimum genes, 3% genes. For statistical analysis, two-sided hypergeometric distribution tests were performed with Benjamini and Hochberg false discovery rate (FDR)-correction at $<$0.05 significance level.

The gene-term enrichment analysis was performed with Gene Ontology Enrichment Analysis Software Toolkit (GOEAST) released 2016-07-15 Version 11.1. The software calculated the fold enrichment based on the expected value. The expected value is the number of proteins expected in a particular GO category, based on the reference list available in the database. Subsequently, it calculates the fold enrichment based on the expected values for a particular GO term, and the probability is determined using a binomial statistic for p-values with a minimum cutoff of 0.05.

Transcriptional Analysis at mRNA Level by qRT-PCR

The qRT-PCR analysis was performed, as explained earlier (Ali et al., 2017). Cells were grown at 80% confluence, three times cold PBS washed, mild trypsinized and pelleted by centrifugation at 400 $\times$ g for 2 min and samples were collected for RNA isolation. RNA was extracted using the RNAeasy mini kit (Qiagen, Germany) followed by cDNA synthesis using cDNA Synthesis kit (MBI, Fermentas). The reactions were carried out for 25 μl total volume containing 1 μl template, 10 nM of 1.25 μl primers, 12.5 μl master mix (Thermo Fisher Scientific, United States) and 9 μl nuclease-free water keeping the standard conditions in 35 PCR cycle (Ali et al., 2017). The associated fold change was measured using the $2-\Delta\Delta CT$ method. The standard t-test statistic was applied for the comparison of the expression of the gene ($p \geq 0.05$). Data were analyzed using MS Excel 2007 and Prism software 5.01 (GraphPad Software, United States).

Cell Viability Tests: MTT; BrdU; Caspase Activity Assay for Apoptosis Determination

To assess the cell viability of MFG-E8 KD and WT cells, we used four different principles based on apoptosis determination assay. For all the assays, cells were initially seeded in a cell count of 1 $\times$ 10^5 cells/ml in a 96-well plate with different intervals of time (12, 24, 36, 48, 60, and 72 h) including respective controls containing complete medium without cells. For MTT-assay [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma)] MTT solution of 0.5 mg/ml dissolved in PBS was added in accordance to the mentioned time intervals and incubated for 2 h at 37°C in 5% CO_2 followed by addition of the DMSO to dissolve the precipitated formazan and absorbances were measured at 570 nm using NanoQuant plate reader (TECAN, Infinite 200 PRO, Mannedorf, Switzerland).

[1]http://www.genome.jp/kegg/mapper.html

The BrdU (5-bromo-2′-deoxyuridine) assay was performed according to the manufacturer's protocol using the assay kit (Cat No. Q11A58, Calbiochem, United States). As mentioned above, cells were seeded in a 96-well plate but with 100 μl medium containing the density of 200 cells per well. The BrdU labels were incubated for 2 h in the growing cells (KD and WT transfected) for integration, followed by measuring the absorbance at 540 nm using the Nano-quant plate reader.

To measure the degree active Caspase- 3/7, we used Caspase-Glo 3/7 activity assay (Promega Cat No. G8091, Madison, WI, United States) according to the manufacturer's protocol with minor changes as also described previously (Kaur et al., 2017). Briefly, cells were seeded and incubated as mentioned above in 96-well plate, and after completion of the individual incubation period, 100 μL of Caspase-Glo 3/7 reagent was added. Further, the plate was incubated for 3 h at room temperature, and luminescence intensities were measured using excitation wavelength 492 nm and emission wavelength 535 nm.

Cell Confluence, Wound Healing, and Transwell Assay

The number of 4×10^4 KD, WT, KD + HA 14-1, and WT + HA 14-1 BuMEC cells were cultured in 24 well cell culture plates to ∼10% of confluence. Cell proliferation was monitored every 7 h using Bio-Rad TC20 automated cell counting system (Bio-Rad, CA, United States) up to 56 h for over 9-time points.

The migratory behavior of KD and WT BuMEC cells was evaluated using the wound healing scratch assay. An equal number of cells were seeded on 6-well plates in six biological replicates and allowed to reach 90% confluency. Upon reaching the suitable confluent stage the cells were then 12 h serum-starved followed by creating the wound by spanning the line at the bottom of the individual wells by using a sterile p-20 pipette tip. All the wells were rinsed gently with DPBS and then one set of the KD and WT cells were treated with HA 14-1, a well-known apoptosis inducer. As mock control cells were also preincubated with DMSO alone. The quantification was manually completed in the scratch cross-sectional area by measuring the number of cells migration.

In the transwell assay, using transwell buckets (Millipore, United States) was covered to the bottom of the inner wells. A cell suspension (5×10^4 cells in 200 μL of DMEM without FBS) was added to the upper chamber, and DMEM with 10% FBS (500 μL) was added to the lower chamber. After 15 h, the migrated cells were stained with hematoxylin. The stained cells were analyzed by microscope in five randomly selected fields/assay. The Image J software counted the cell numbers. Values represent the means ± SD from six replicates.

In vitro Phagocytic Assay

The phagocytosis assay was performed using BoMAC and the BuMEC cell line (KD and WT) with the help of a previously published protocol (Hanayama et al., 2002). The apoptosis was induced by HA 14-1 in the assay. The respective BuMEC (1×10^6 Cells) KD or WT with or without apoptotic inducer were added to 2×10^4 macrophage cells previously cultured in eight well Lab-tec II chamber slides and incubated for 1.5 h followed by the fixation using 1% paraformaldehyde and TUNNEL staining was performed. The TUNNEL positive apoptotic cells per phagocytes were counted (phagocytosis index) by light microscopy in a total of 200 phagocytes, and the relative index was determined.

Statistical Analysis

These statistical calculations were performed using the GraphPad Prism 4 software package (GraphPad Software, Inc., La Jolla, CA, United States). Data were shown in the form of bar plot as mean with standard error of mean (±SEM) with relative 100 percent to control. Statistical significance for MTT, BrdU, Caspase 3/7, and cell grown confluence assay was determined using a two-way ANOVA test with Bonferroni post-test correction for FDR-correction at <0.05 significance level indicated by asterisk marks. For wound healing, transwell, and phagocytosis assay one-way ANOVA test with repeated measures along with post Turkey test for FDR-correction at <0.05 significance level indicated by asterisk marks.

For mass spectrometer and qRT-PCR results, data are represented as relative $\log_2$ fold change. Transcripts CT values were normalized against the housekeeping gene β-actin, and fold change was calculated using $2^{-\Delta\,\Delta}$ CT method. All experiments were performed in triplicates. The significance of the differences in expression levels was tested using the non-parametric Mann–Whitney U test $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

RESULTS

Analysis of MFG-E8 Protein Sequence and Evolutionary Relationship

The mammary gland remodeling is a highly programmed, genetically controlled process involving cell death during involution that is both spectacular in scope and well-organized in execution. As shown in multiple reports that MFG-E8 is an essential protein for the successful accomplishment of the involution process (Stubbs et al., 1990; Lönnerdal, 2003; Atabai et al., 2005; Hanayama and Nagata, 2005). Our goal of this study is to determine the associated regulatory proteins and signaling pathways linked with MFG-E8 protein through its knockdown in buffalo MECs. As the information on the role of MFG-E-8 in bovine is scanty, we examined the sequence and its domain organization for evolutionary conservation (accession number NP_001277850.1). The same mfge8 full-length gene sequence used in multiple sequence alignment (MSA) was used for the preparation of shRNA for knockdown experiments (**Supplementary Figure 1**). The MSA of MFG-E8 from 35 organisms showed the absolute conservation of the RGD motif that is involved in cellular interactions with cell-surface integrins (Taylor et al., 1997), but we argued that its presence may be fortuitous and may imply for another cellular process. Likewise, we determined another NPC novel conserved motif among the organisms, not discussed in the previous literature. Nonetheless, the deletion of 42 amino acid residues was observed (28th to 70th aa) in the MFG-E8 sequence of a human, baboon, Ghana monkey, and Bolivian monkey, suggesting its role in evolutionary

selection. Interestingly, in the same region in remaining all organisms, a conserved GGTC sequence was observed. The result implies that the later got diverged during evolution and lost 42 aa length peptide (**Figure 1A**).

Further using full-length protein MSA (**Supplementary Figure 1**) and different databases, we determined the five conserved domains namely: 1 signal peptide, 2 EGF domains (respective e-values are 0.000197 and 0.00000435), and 2 FA58C domains (respective e-values are 6.56e-41 and 2.01e-36) in the protein sequence which were found consistent among all 12 databases and algorithm searched (**Figure 1B**). Rest all other identified domains have been detailed in the **Supplementary Table 2a**. The conservation, quality, and consensus among protein sequences revealed that the residues from 115th to the 185th amino acid at the N-terminal is non-conserved. The results were consistent, which were validated by the four independent algorithms for the calculation of protein disorderliness (**Figure 1C**). The prediction of the secondary structure identified eight α-helix and twenty-two β-sheets, and these results were consistent among three different secondary structure programs (jnetpred, JNETPSSM, and JNETHMM) (**Figure 1D**). Further, the determination of motif recognition analysis using the motif-x algorithm showed the total identification of eighteen motifs corresponding to eight specifics for phosphorylation and ten for glycosylation (for detailed motifs information, see **Supplementary Figure 2**). The phylogenetic analysis of MFG-E8 from 35 organisms showed that buffalo is highly close to bovine, goat, sheep, and deer and comes in one cluster out of eight classified groups. The results were validated with the help of principal component analysis (PCA) analysis using the BLOSUM62 model of the MSA. In the PCA plot, a similar sequence lies close to each other. We identified the highly divergent nature of MFG-E8 in alligator, anoles, chicken, and mallard birds, which were grouped into a separate cluster (**Figures 1E,F**). Altogether, these results suggest that intradomain regions are conserved among the organisms, but many differences were observed in the interdomain areas.

Impact of MFG-E8 Repression on the Morphology and Phagocytic Activity of MEC

Phosphatidylserine-MFG-E8 axis mediated cell death modality is evolutionarily conserved, which has been recently explained (Birge et al., 2016). In support, we identified the conserved MFG-E8 domains in all the species, signifying for the necessary cellular proliferation process and involvement in phagocytosis. The transfection, retrograde passaging, and continuous selection using puromycin antibiotic resulted in stable expression of shRNA. The presence of GFP in the vector helped for the positive selection of transfected clones. The qRT-PCR analysis showed 90-fold down-regulation in the expression of MFG-E8 transcripts (**Figure 2i**) and the same results were validated by western blot analysis. The dense expression of GFP shows the expression of a scramble as well as MFG-E8 specific shRNAs (**Figures 2ii A,B**).

Interestingly, the low expression level of inherent MFG-E8 protein resulted in the formation of globule-shaped structures that are not seen in the healthy or scrambled shRNA transfected cells (**Figures 2ii C,D**). However, the formation of these structures was prominent at 72 h in KD BuMEC cells (**Figures 2ii E,F**). To exclude the possibility that GFP is stimulating these globules aggregation, we transfected non-target scramble shRNA GFP vector to BuMEC cells as a negative control, and no such structures were observed (**Figures 2ii C,D**). To understand and discover the MFG-E8 associated interconnection of pathways in the regulation of MEC proliferation, apoptosis, and transformation, we performed iTRAQ based comprehensive quantitative proteomic analysis between KD_MEC and WT_MEC and validated the findings through multiple *in vitro* assays as discussed later.

As MFG-E8 was previously described as a bridging molecule between macrophages (MQ) and apoptotic cells (Oshima et al., 2002), we determined the phagocytic index of MFG-E8 knocked down MEC cells. For the control, we performed the assay alone for Bu (Normal or wild type BuMEC cells), Bu(+) (Apoptotic induced BuMEC), Bu(#) (scrambled shRNA-MFG-E8 BuMEC cells), Bu(*) (MFG-E8 Knockdown BuMEC cells), and MQ (BOMAC) cell. Surprisingly, apoptotic induced BuMEC cells showed the phagocytosis behavior but not others. For assay we tested two separate control one with scramble shRNA and other with WT cells. The results showed the minimal phagocytic index in Bu Bu(*) and Bu Bu(#) (0.42) than Bu(*) MQ (0.49) and lastly to Bu(#) MQ and Bu MQ (0.73) (**Figure 2iii**). The suggested reasons for low phagocytosis index in MFG-E8 knocked down Bu(*) cells may be the low expression of MFG-E8 or the decreased level of its associated receptors. Later the same results got confirmed in the proteomics data.

On the other hand, the maximal phagocytosis index was seen in the apoptosis induced Bu(+) MQ (1.81), which was used as a positive control. Unexpectedly, we observed the higher index in apoptosis caused Bu(+) Bu(*) and Bu(#) Bu(*), co-cultured (1.50) and in apoptotic induced Bu Bu(+) and Bu(#) Bu(+) (1.21). The detailed information about the ANOVA results has been provided in the **Supplementary Table 2b**. The phagocytosis assay results support previous findings that the healthy epithelial cells act as amateur phagocytes and are involved in the clearance of apoptotic cells through the expression of MFG-E8 (Monks et al., 2005, 2008).

Further, we examined the cell growth percentage in KD and WT cells up to 56 h; the results confirmed the significant high growth disparity in WT cells $p < 0.001$ at 56 h in relation to MFG-E8 KD. Subsequently, we tested the growth rate in the presence of the apoptotic inducer HA14-1, it resulted in the reduction of the KD cell population at 56 h to 52% growth, and WT treated with HA14-1 at 56 h to 55% growth compared to control with significant difference p-value <0.001. In comparison to wild type, the initial significant difference for all three conditions was observed from 21 h of the cell growth assay (**Figure 2iv**). Next, we performed the wound healing assay using the same four conditions. The analysis of variance showed that in the first comparison of WT versus WT + HA14-1 mean difference-20.33 ($P <$ value 0.0001), in second comparison WT versus KD, mean difference-15.00 ($P <$ value 0.0001), and in the third comparison WT versus KD + HA14-1 mean difference-31.17

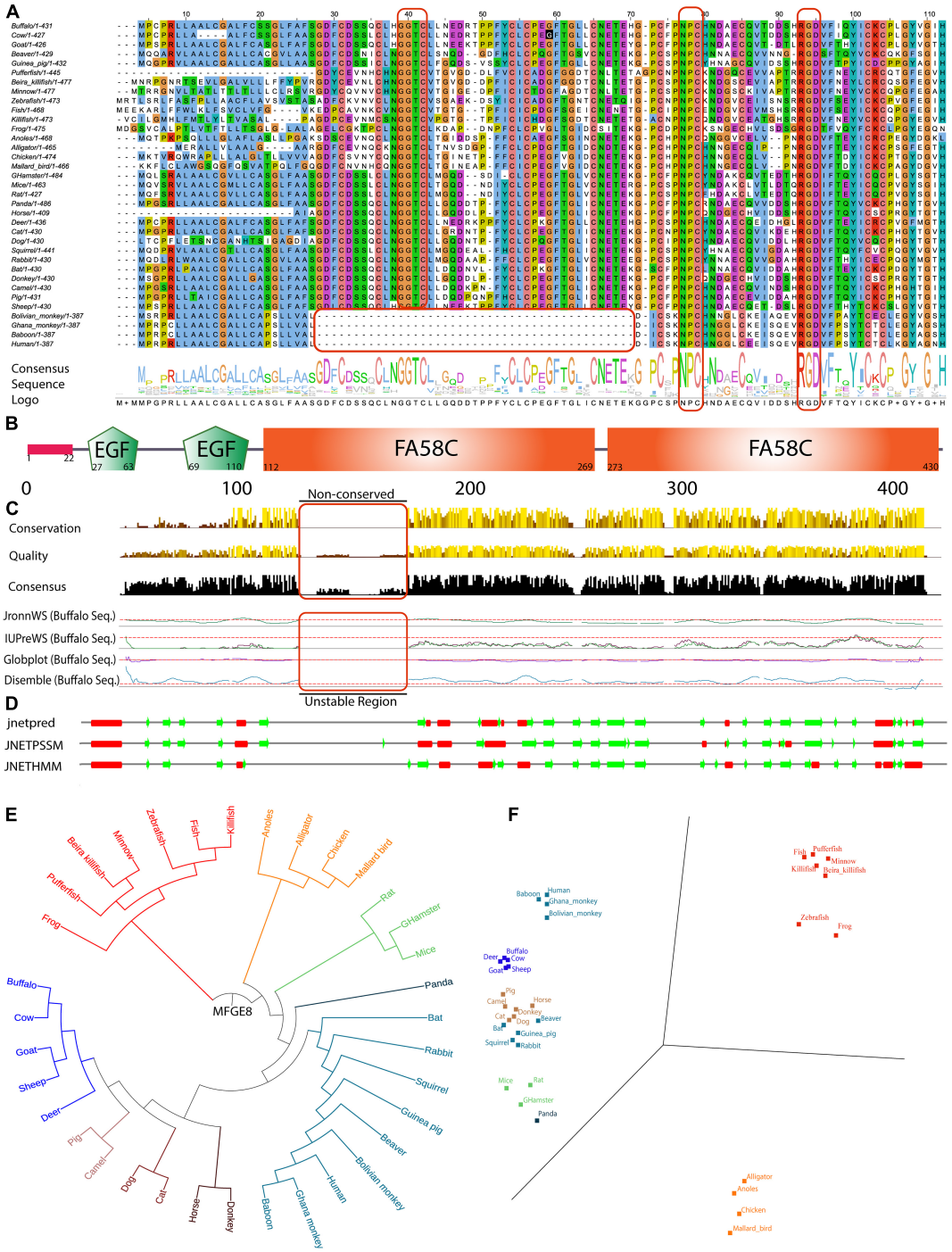


FIGURE 1 | *In silico* based MFGE8 protein sequence characterization. **(A)** Multiple sequence alignment (MSA) of 35 organisms used from UniProt non-redundant and reviewed database. Total numbers of amino acid residue present in the individual sequence were reported in front of the organism name, and MSA is clustal color-coded. The consensus sequence logo was shown in the downside of the alignment, and the conserved nature was depicted using the size of the amino acid single alphabet symbol. **(B)** The structure of the MFGE8 protein containing the complete common domains observed in all programs. **(C)** The full-length sequence quality, conservation, and consensus nature of MSA are shown in comparison with 35 organism-specific MFGE8 protein. The buffalo specific MFGE8 protein sequence disordered behavior was analyzed using four algorithms (JronnWS, IUPreWS, Globplot, and Disemble), keeping the parameter defaults for all of them. The graph showed consistent results among the programs. **(D)** The secondary structure was predicted for a full-length MFGE8 sequence with the help of three independent programs (jnetpred, JNETPSSM, and JNETHMM). The red mark shows the α-helical structures, while the green arrows show the β-sheets structures in the MFGE8 protein. **(E)** Phylogenetic tree analysis of all 35 organisms used using distance tree. **(F)** Principal component analysis of MSA for all 35 organisms using BLOSUM 62 algorithm.

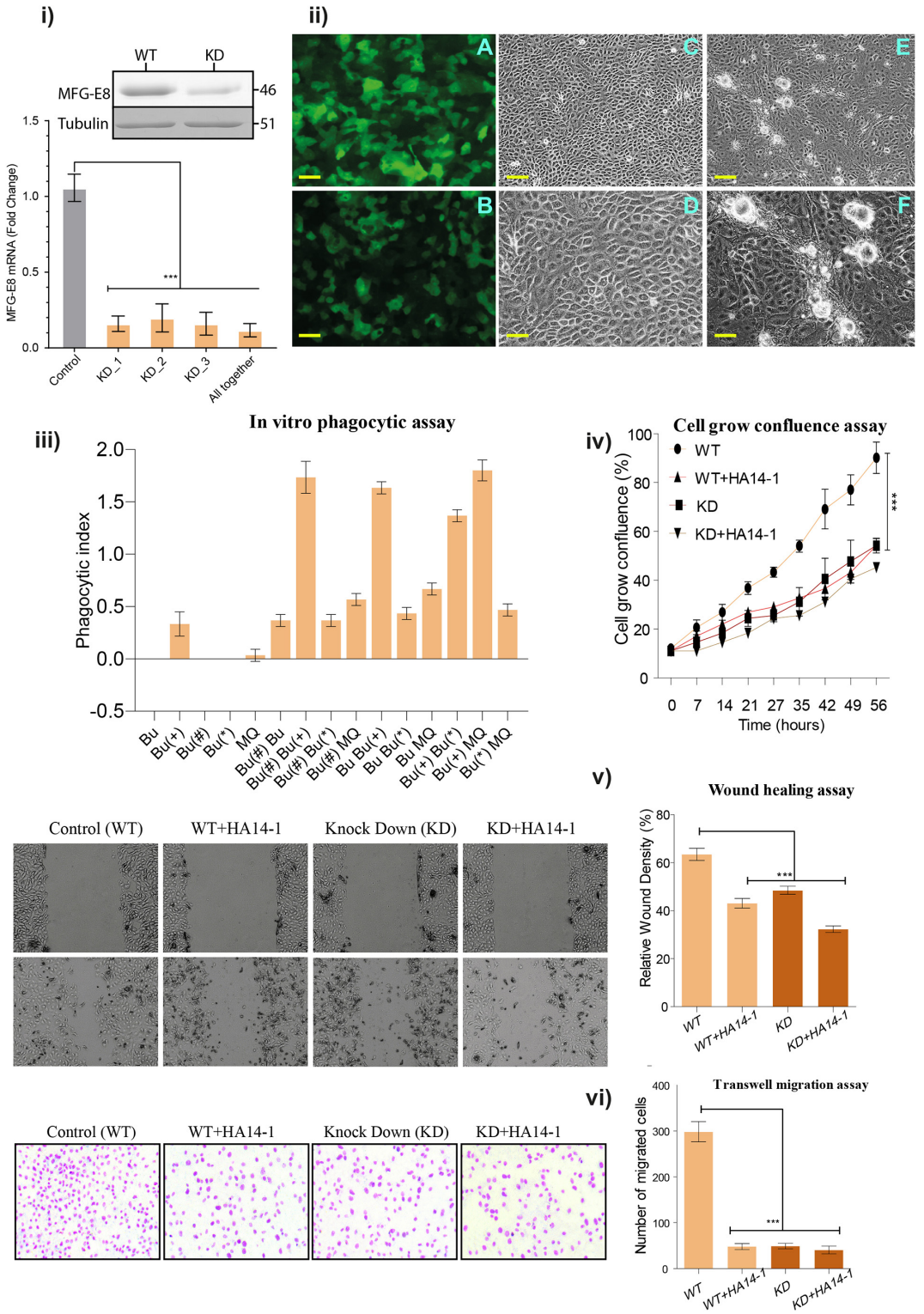


FIGURE 2 | (i) Quantitative real-time PCR assay: Real-time PCR reaction was performed using three independent shRNA (KD_1, KD_2, and KD_3) and in a combination of all together. All the reactions were performed in three biological replicates and repeated at three replicates, and values are shown in ± SEM as error bars. **(ii)** Stable transfection of buffalo mammary epithelial cells (BuMECs) cells. **(A)** Transfected BuMEC cells over-expressing non-target scramble shRNA pLKO.1-puro-CMV-tGFP vector. **(B)** Transfected BuMEC cells with MFG-E8 specific shRNA containing pLKO.1-puro-CMV-tGFP vector, respectively. **(C)** Transfected BuMEC cells over-expressing non-target scramble shRNA pLKO.1-puro-CMV-tGFP vector at 72 h of growth, magnification 200×. **(D)** The same panel-C cells at

(Continued)

FIGURE 2 | Continued
higher magnification 600×. **(E)** Transfected BuMEC cells with MFG-E8 specific shRNA containing pLKO.1-puro-CMV-tGFP vector at 72 h of growth, magnification 200×. **(F)** The same panel-E cells at higher magnification 600×. Scale bar is 100 μm. **(iii)** *In vitro* phagocytic assay. The co-culture phagocytic assay was performed to determine the phagocytosis ability in the absence of MFG-E8 and the induction of apoptotic inducer (HA 14-1). The graph axis is denoted with Bu (normal or wild type BuMEC cells), Bu(+) (apoptotic induced BuMEC), Bu(#) (scrambled shRNA-MFG-E8 BuMEC cells), Bu(*) (MFG-E8 knockdown BuMEC cells), and MQ (BOMAC). **(iv)** Cell grown confluence assay. The WT and KD cells were used along within the presence or absence of HA14-1 inhibitors. The cells were grown with 7 h time interval for 8-time points up to 56 h. The experiments were performed in three independent biological replicates. **(v)** Wound healing assay. The same four conditions were used for the wound healing assay. All the reactions were performed in three biological replicates and repeated at three replicates, and values are shown in ± SEM as error bars. **(vi)** Transwell migration assay. The migration behavior of BuMEC upon knockdown and normal condition along with and without HA 14-1 inhibitor after 48 h was detected by wound healing assay. Cell migration was calculated as described in Section "Materials and Methods." Values are normalized to their respective 0 h time point and presented as means ± SEM of three independent experiments. All the reactions were performed in three biological replicates and repeated at three replicates. Values with $***p < 0.001$ are statistically significant. ns, non-significant.

($P <$ value 0.0001) (**Figure 2v**). The results describe that KD or treatment with the HA 14-1 inhibitor showed a remarkable reduction in the proliferation ability during would healing assay. Comparable results were perceived in the transwell migration assay with a significant difference with WT cells for cellular migration (**Figure 2vi**).

Widespread Proteome Changes Triggered by Ablation of MFG-E8

We employed the loss of function strategy and optimized the iTRAQ based global quantitative proteomics workflow to unravel the regulators of MFG-E8 mediated signaling pathways crosstalk. The protein samples were prepared from 5 days grown MFG-E8 deficient and wild type epithelial cells, followed by digestion, iTRAQ labeling, fractionation, and LC-MS/MS-based identification and quantitation. The combined workflow trans proteomics pipeline analysis enabled the identification of 4,78,552 total spectra using three search engines, which represented 12,649 proteins having an iProphet probability of ≥0.999 and error rate <0.00001 (**Figures 3A,B**). The Libra based relative quantification of non-redundant protein entries resulted in 8858 DEPs (p-value < 0.01). Only those proteins were selected for analysis, which qualified the inclusion criteria of ≥2 peptides, minimum cutoff MS/MS count 2, minimum of seven amino acid residues, and iProphet probability ≥ 0.95 per peptide (**Figure 3C**). Subsequently, the precision in the data acquisition was determined using statistical correlation among technical and biological replicates, which have been shown as multiple densities scatter plots (**Supplementary Figure 3**). The Pearson correlation R^2 value of 0.985 and 0.986 was observed between the biological replicates of KD and WT cells, respectively (**Supplementary Figure 4**). The plots show the precision in acquired MS/MS spectral intensities used for the quantification. A list of identified proteins, iProphet protein probability, sequence coverage, peptide sequence, Libra quantification values, and fold change values between MFG-E8_KD and Wild type control cells has been provided in **Supplementary Table 1**. Proteomics findings were verified with results obtained from various cellular assays (**Figure 3D**).

The profiled ultra-deep MEC proteome was cataloged, and individual proteins were predicted for its localization in the cellular compartments based on the consensus motifs available in the full-length protein sequences. Remarkably, we identified 45% of the whole proteome to be associated with the nucleus, while 20% of proteins were from the cytoplasm. With the help of our optimized proteomics workflow, including the microwave energy pretreatment of the proteins before digestion for optimal cell lysis, it allowed us to identify 17% of plasma membrane protein without any additional enrichment step. The remaining 18% were distributed in the other cellular compartment including peroxisome 0.4%, cytoskeleton 3%, endoplasmic reticulum 3%, Golgi 1%, lysosome 1%, junction proteins 0.9%, mitochondrial 3%, and extracellular or secreted proteins 6% (**Figure 4A**). The downstream statistical analysis in the R- environment using a two-sample t-test based modified SAM-test showed significant proteins in the volcano plot (**Figure 4B**). For all DE proteins corrected p-values were calculated assuming the equal variance in biological replicates while keeping the *fudge factor* s0 = 1 and FDR = 0.01, as the constant parameters to locate out the "t-test Significance," "−Log t-test p-value," "t-test Difference" and "test statistic" for all entries. The volcano plot provided an estimate for both sided significance values of proteins (**Supplementary Table 1**). Our statistically defined data portrayed that the knockdown of MFG-E8 changes the dynamics of proteome expression in MECs, indicating its crucial role in basic cell proliferation.

The majority of highly down-regulated proteins in KD_MEC cells included RFX8, SIX6, ELP3, ZMYM3, EIF4, ENIF1, MYH14, DNAH11, TSBP, CBLN4, CNP, OR14C36, and CD2BP. They belong to protein classes of cytoskeletal protein (PC00085), transferase (PC00220), and transcription factor (PC00218), hydrolase (PC00121), an enzyme modulator (PC00095) (**Supplementary Table 3**). The up-regulated proteins in KD_MEC cells included LYVE1, FOLH1, PRKCSH, TRPV3, RING1, DOCK1, PLTP, SLC16A13, LENG1, and DMTN which are involved in the functions of transporter activity (GO:0005215), receptor activity (GO:0004872), signal transducer activity (GO:0004871), catalytic activity (GO:0003824) (**Supplementary Table 4**). The proteins perturbed by the suppression of MFG-E8 expression in the epithelial cell showed the significant rearrangement of 572 membrane trafficking, 220 ion channel-specific, and 202 cytoskeleton proteins in KEGG annotation (**Supplementary Figure 5** and **Supplementary Table 5**). The variety of ion channels include calcium-activated potassium channel, potassium voltage-gated channel, potassium channel regulatory protein, epithelial chloride channel, sodium channel protein, amiloride-sensitive sodium channel, transmembrane channel-like protein, short

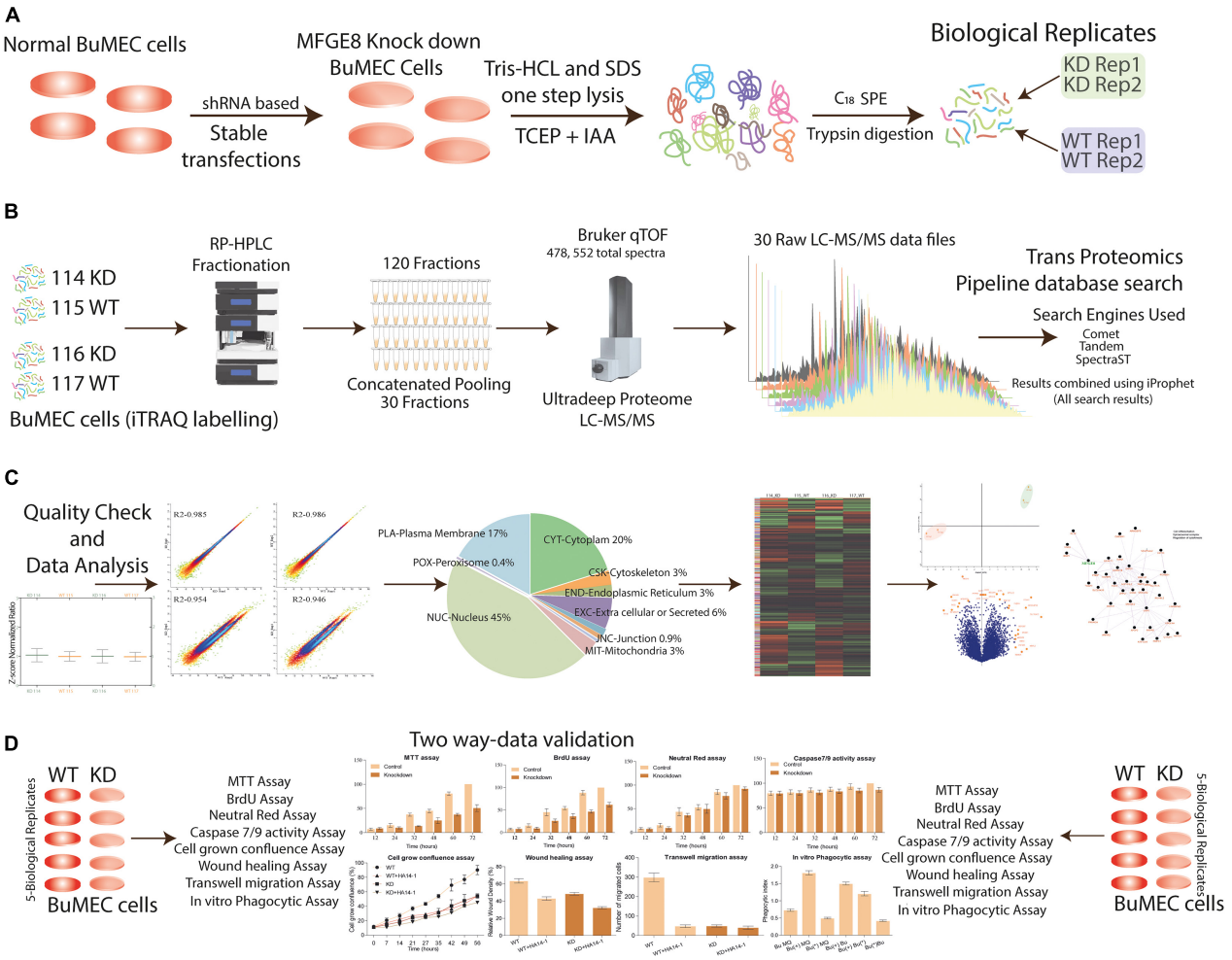


FIGURE 3 | Experimental design and quantitative proteomics workflow allow a comprehensive analysis of MFGE8 affected proteome. **(A)** BuMEC cells were transfected with shRNA for stable knockdown of MFGE8 protein expression, stably transfected cells grown for 120 h followed by quick lysis, protein extraction, quantification, and trypsin digestion for peptides generation in two biological duplicates (see the section "Materials and Methods"). **(B)** Peptides were labeled with iTRAQ labels, which are 114, 116 for KD (knockdown cells), and 115, 117 for WT (wild type cells treated with scrambled shRNA) followed by reversed-phase UHPLC based peptide fractionation. The fractions were concatenated pooled into 30 final fractions, and iTRAQ labeled peptides were analyzed by high-resolution qTOF-LC-MS/MS. The database search analysis was performed by using three search engines (Comet, Tandem, and SpectraST) in the environment of the transproteomics pipeline, further, the description of raw data processing as described in Section "Materials and Methods". **(C)** Outline of mass spectrometer data processing and analysis is shown; necessary detailed methodologies are provided in the Section "Materials and Methods." For relative quantification, data were filtered and normalized to make sure fair evaluation across all duplicate biological samples. The resulting identified proteins were used for their localization in the cell compartments. The heat map was generated for clustering the group of proteins. Principal component analysis was performed to verify the isolation of the components of different labels. Finally, the volcano plot was graphed to identify the significantly differentially express proteins. These DEPs were used for network analysis, and lastly, the biological MFGE8 based regulatory pathway was predicted. **(D)** The validation using the different biological assays.

transient receptor potential channel 1, transient receptor potential cation channel, voltage-dependent anion-selective channel, and ATP-sensitive inward rectifier potassium channel (**Supplementary Table 1**). Another interesting observation is the up-regulation of the dedicator of cytokinesis protein isoforms; DOCK1, DOCK2, DOCK3, DOCK4, DOCK6, DOCK8, DOCK9, DOCK10, and DOCK11. The cytoskeletal changes and dramatic upregulation of transporter activity proteins together probably contribute to the distortion in the shape of KD_MEC cells on the fifth day of the culture, as seen in **Figure 2ii**. The size of the MFG-E8 knocked down cells as measured in terms of area and perimeter over the different points of observation, i.e., 12, 72, and 120 h of cell growth showed significant expansion (p-value ≤ 0.001). It may be associated with ionic imbalance inside MECs due to perturbation in many membrane channel proteins.

Our previous results of LIF overexpression also showed the activation of these identified membrane channels proteins, which resulted in the formation of globule structures in cells (**Figure 2ii** and **Supplementary Table 1**) (Kaur et al., 2017; Ali et al., 2018). We assessed all DEPs using the PCA plot to gain insight for its distribution in component 1 (49.2%) and component 2 (31.8%) with the identification of orthogonally transformed of

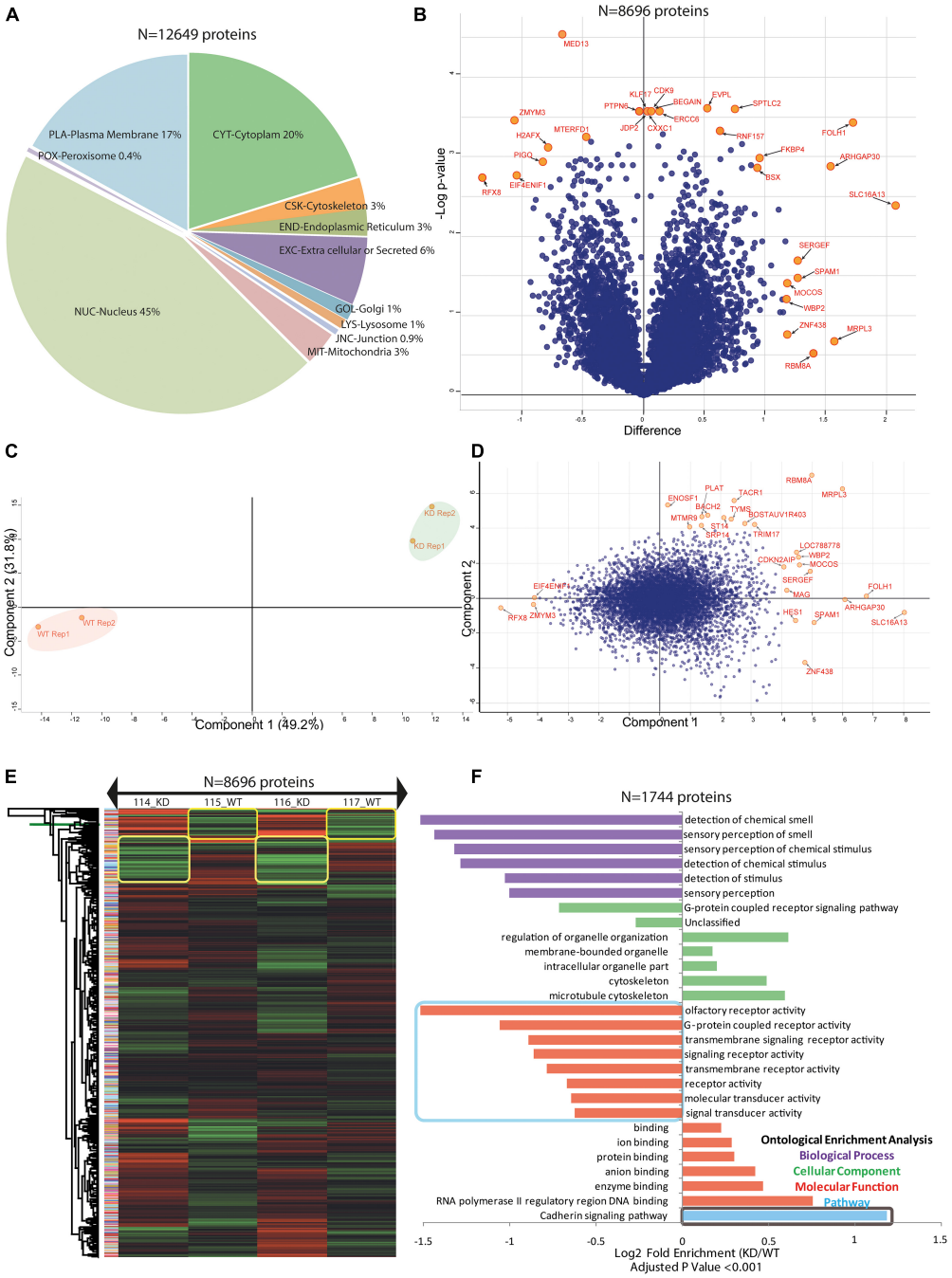


FIGURE 4 | BuMEC in-depth proteome analysis. **(A)** A pie chart shows the different percentage of occupancy of the identified proteome in compartmental based localization of epithelial cells. See also **Supplementary Table 5**. **(B)** Volcano plot shows the significantly differentially expressed proteins and extremely affected proteins were labeled with the orange color circle with named in red color. The two-sided *t*-test was calculated using artificial within-group variance S0 equals zero. Further truncation was performed using a permutation based FDR equals 0.01, and the number of randomizations equals 250. The *X*-axis is the *t*-test difference values while the *Y*-axis is –log of adjusted *p*-values. For detailed data, see also **Supplementary Table 1**. **(C)** Principal component analysis shows the proteome data isolation of the components of different iTRAQ labels for biological duplicates samples. Highlighted color reflects the replicate groups; the red color is for wild type cells (WT), and green color is for knockdown cells (KD). **(D)** The biplot of PCA shows proteome data for comparison of component 1 (*X*-axis) to component 2 (*Y*-axis) for WT and KD cells in differentially identified proteins. Blue dots that are closely associated are highly correlated in terms of the observed proteome, and extremely isolated orange circles are separated proteins from PC1 and PC2 and were labeled as red color. **(E)** Heat map presentation of a hierarchical cluster of all 8696 proteins that show significantly different ($p \leq 0.05$) relative abundances in WT versus KD cells. The green color depicts low while red color depicts high expression levels. **(F)** Protein ontologies enrichment analysis is shown in the form of a bar graph. Biological process (BP), cellular component (CC), molecular functions (MF), and pathways were color-coded with bars. All the enriched values were $\log_2$ transformed, and the adjusted *p*-values used as <0.001. For all groups, the overall number of fold enrichment is shown in comparison to the reference database, and detailed information is provided in **Supplementary Table 1**.

essential proteins, indicated in the red label (**Figures 4C,D**). The majority of the proteins identified as PCA outliers showed their involvement in the cell cycle regulation in consonance with GO analysis and KEGG annotation (**Supplementary Tables 3, 4, 5**, respectively). The comparison using hierarchical cluster analysis with Euclidean distance and average linkage with K-mean 300 clusters revealed significant changes in the expression level of all 8858 proteins (**Figure 4E**) which is same as shown in the volcano plot (**Figure 4B** and **Supplementary Table 1**; 4,025 up-regulated genes, 4,833 down-regulated genes, false-discovery rate-adjusted $P < 0.05$). Thus, our result strongly suggests that knockdown of MFG-E8 protein directs widespread cellular response that is reflected by the changes in the whole proteome profile of MEC, requiring additional investigation into prominent cellular pathways and linked proteins that can distinguish the assault control mechanism from the direct downstream effects of MFG-E8 knockdown.

We observed that the suppression of MFG-E8 expression increased the duration of the G0-G1 phase transition by 32.62% in comparison to WT_MEC 18.4%. However, the S-phase duration is shortened in KD_MEC from 65.85% to 47.20% (**Supplementary Figure 6**). To explain this effect, we looked at the Gene Ontology classification of DEPs by GO analysis (see Materials and Methods section). The investigation revealed significant enrichment of 530 GO terms with adjusted *P*-value < 0.05, including BP (321), CC (94), MF (86), Reactome/pathways (11), and protein class (18) terms, respectively (**Supplementary Table 2**). We further scrutinized the main GO terms simultaneously using fold change cut off (2.0-fold up and 0.5-fold down-regulated) and highly significant *P*-value < 0.01 in each category to visualize the pervasive impact of low MFG-E8 level in cells (**Figure 4F**). Surprisingly, we found the over-enrichment of the process for the negative regulation of intracellular signal transduction (GO: 1902532) and negative control of the cell cycle (GO: 0045786). It supports our observation of an increase in the cell cycle, doubling the time of KD_MEC cells (**Supplementary Figure 6**). It is important to note here that the over-representation of key GO terms significantly enriched in the biological process showed posttranscriptional gene silencing (GO: 0016441) and gene silencing by RNA (GO: 0031047) which proves the shRNA mediated MFG-E8 silencing in our experiment. Further exploration of results showed that MFG-E8 knocked down disturbs the regulation of Ras protein signal transduction (GO: 0046578), cell-substrate adhesion (GO: 0031589), extracellular matrix organization (GO: 0030198), DNA repair (GO: 0006281), protein autophosphorylation (GO: 0046777). The under-representation analysis indicated the predominant effect on nucleosome assembly (GO: 0006334), G-protein coupled receptor signaling pathway (GO: 0007186), natural killer cell activation involved in immune response (GO: 0002323), detection of stimulus (GO: 0051606), Fc gamma R-mediated phagocytosis (bta04666) providing the evidence for a low phagocytic index of KD cells.

We mapped DE proteome ($P < 0.05$) using genemania, reactome, and string database in a single platform on Cytoscape into different cellular processes to illustrate interdependent proteins regulation under the signaling of MFG-E8 (**Figure 5**). Overall, with these results, we understand that the repression of the MFG-E8 affects a wide variety of proteins in MEC leading to a severe breach in cellular homeostasis.

MFG-E8 Associated Proteome Regulates MEC Homeostasis Through Cytoskeleton Rearrangement

To understand the process of homeostasis in the absence of MFG-E8, we executed four powerful databases; Gene Ontology Enrichment Analysis Software Toolkit (GOEAST), ClueGO, DAVID, and KEEG (**Figure 6**). The over-enrichment of the prominent cellular component containing terms included GO:0005737 cytoplasms (12.35), GO:0016020 membrane (10.86), GO:0005654 nucleoplasm (9.06), GO:0005871 kinesin complex (7.77), GO:0005622 intracellular (6.72), GO:0005813 centrosomes (6.51), GO:0005925 focal adhesion (5.64), GO:0005829 cytosols (5.44), and GO:0008270 Ion channels (6.21). The **Supplementary Tables 2, 3** enlist the details of all other enrichment. This analysis highlighted the negative regulation of cytoskeleton organization (GO:0051494), microtubule cytoskeleton organization (GO:0000226), which was in parallel to the finding of morphological changes in KD_MEC cells (**Figure 2ii**). The globular structure formation as a result of cytoskeletal perturbances has been reported in other studies as well, where stat3 mediated Na-K + channels are the key players (Ali et al., 2017). To gain insight on various pathways and their interconnections, we identified (**Supplementary Figure 5**) bta04510:Focal adhesion (10.19), bta04512:ECM-receptor interaction (10.18), bta04810:Regulation of actin cytoskeleton (10.18), bta02010:ABC transporters (5.90), bta04110:Cell cycle (5.18), bta04390:Hippo signaling pathway (4.67), bta00562:Inositol phosphate metabolism (4.31), bta04070:Phosphatidylinositol signaling system (4.15), bta04010:MAPK signaling pathway (3.97), bta04150:mTOR signaling pathway (3.72). The down-regulation of MFG-E8 thus proves that it is a central molecule in the maintenance of cellular homeostasis. A large number of cellular processes are interconnected through MFG-E8, among which structural components involved in focal adhesion, ECM receptor interaction, and actin cytoskeleton regulation are worst affected in its absence (**Figures 6A,B**) (detailed shown in the **Supplementary Tables 2, 3**).

Immunoregulatory Effects of MFG-E8 Downregulation in BuMECs

We next set forth to determine the immunological behavior of the MFG-E8 associated protein, we mapped the proteome against the immune pathways and identified the B cells/T cells activation, lymphocytes activation, NK cell-mediated immunity, and leukocytes differentiation (**Figure 6C**). In correspondence with the multiple previous reports where MFG-E8 has been shown to elicit the immune response in terms of phagocytic activity in the MECs (Monks et al., 2002; Abrahams et al., 2004; Watson, 2006a,b; Stein et al., 2007; Watson and Kreuzaler, 2011). The detailed information about the identified GO terms is

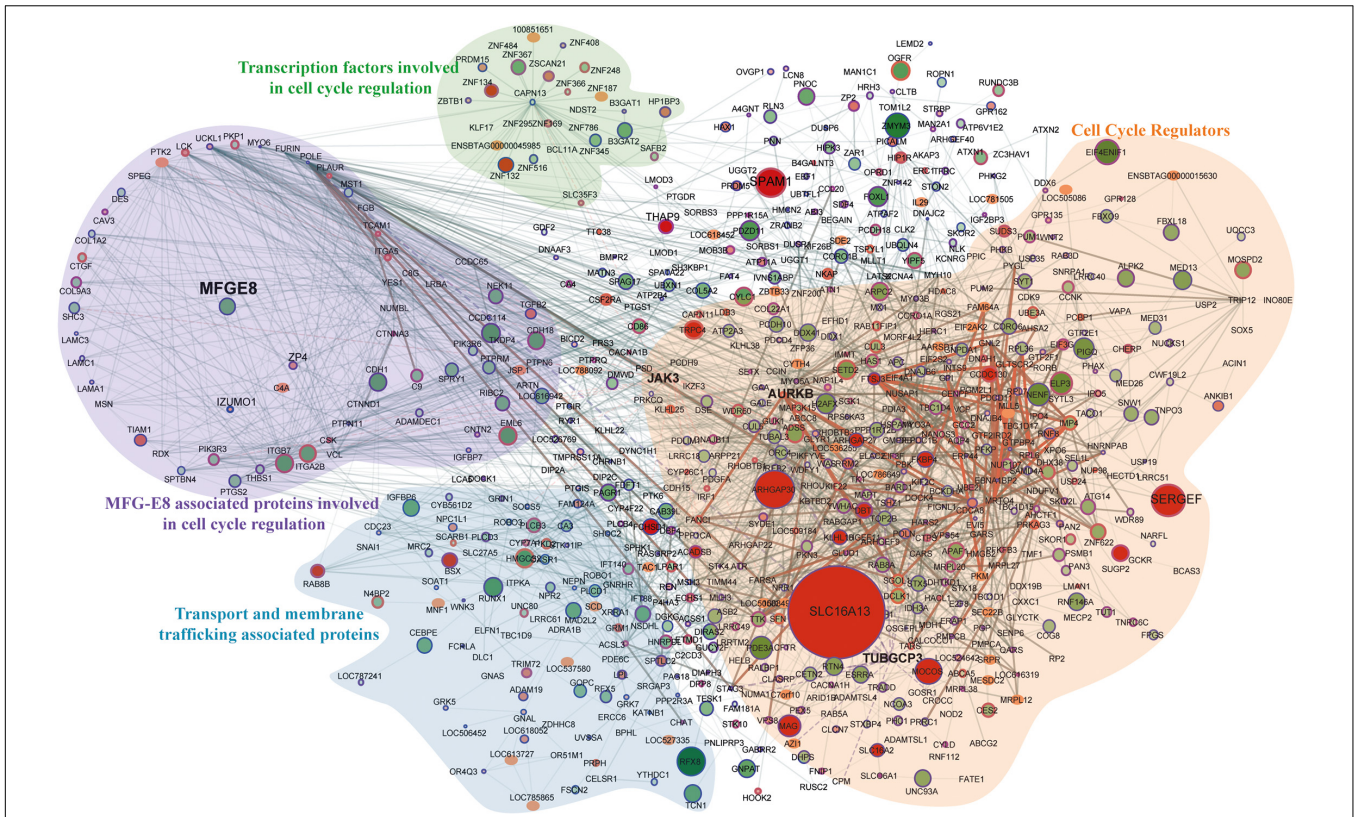


FIGURE 5 | Protein-protein interaction map for top significantly (p-value < 0.001) expressed proteins. The total number of nodes is 682 with 2140 interacting edges, node color and size both show the up/downregulation, specifically, darker color and bigger size node defines the high value of the fold change. The smaller size node defines no observed fold change in proteomics data. The border of the nodes shows the significance of the lesser p-values < 0.001; thicker the border lowest is the assigned p-values. The thickness of the edges describes the experimentally identified co-expressed proteins in the database.

mentioned in the **Supplementary Table 5** with fold enrichment and associated p-values and FDR.

Although MFG-E8 associated genome-wide studies were previously performed but only at the transcriptomics level (Sugano et al., 2011) and as per our knowledge, no MFG-E8 associated proteome-wide global ultra-deep proteomics study has been so far done. Moreover, looking into the importance of hundreds of genes target previously recognized (Sugano et al., 2011), and protein targets identified in our current study (**Figure 4B** and **Supplementary Figure 5**), point out several questions in regards to its diverse downstream associated signaling which remain unresolved. The critical prerequisite of MFG-E8 in basic organ developmental biology; such as mammary gland (Hu et al., 2009; Sugano et al., 2011), intestinal epithelium (Zhao et al., 2012), and lungs remodeling process (Aziz et al., 2015) further enforced to accomplish the comprehensive proteome analysis.

Interestingly, recent findings described that MFG-E8 derived short peptide, known as MSP68, showed immunomodulatory function in sepsis-injured tissue. These studies indicate the unique therapeutic property of MFG-E8 and its derived peptide for regulating the inhibition of the infiltrating neutrophils migration, especially in lungs sepsis (Hendricks et al., 2017; Hirano et al., 2017). In correspondence to these previous results, our proteomics data proves that the MFG-E8 mediated signals toward neutrophils activation and migration are severely compromised. The under enrichment of the immune-related proteins upon MFG-E8 knockdown and provide mechanistic insight into its downstream pathway.

MFG-E8 Affects Signature Transcription Factors (TFs)

Around 45% of the total profiled proteins belonged to the nucleus (**Figure 4A**). We determined the MFG-E8-dependent signature TFs with a cutoff of 1.5-fold up/down-regulated differentially expressed MEC proteome (n = 938 proteins) with significant p-value < 0.001) using a computational regulatory analysis with iRegulon (Janky et al., 2014). The analysis resulted in the identification of 12 TFs, which regulated 806 DEPs out of 938 identified proteins in the nucleus, as shown in the form of the interaction map (**Figure 7A**). The identified TFs and their targets includes ZNF652 (232), CHD2 (222), MEF2A (243), GATA2 (256), FOXF2 (73), HDAC2 (77), FOXA1 (112), JUN (60), TFDP3 (140), POU2F1 (221), NFYA (68), and SNAI2 (23). We identified a common motif of C/ACAATXXXXGCG on the target genome for binding with these TFs, as revealed by the analysis on the JASPAR database

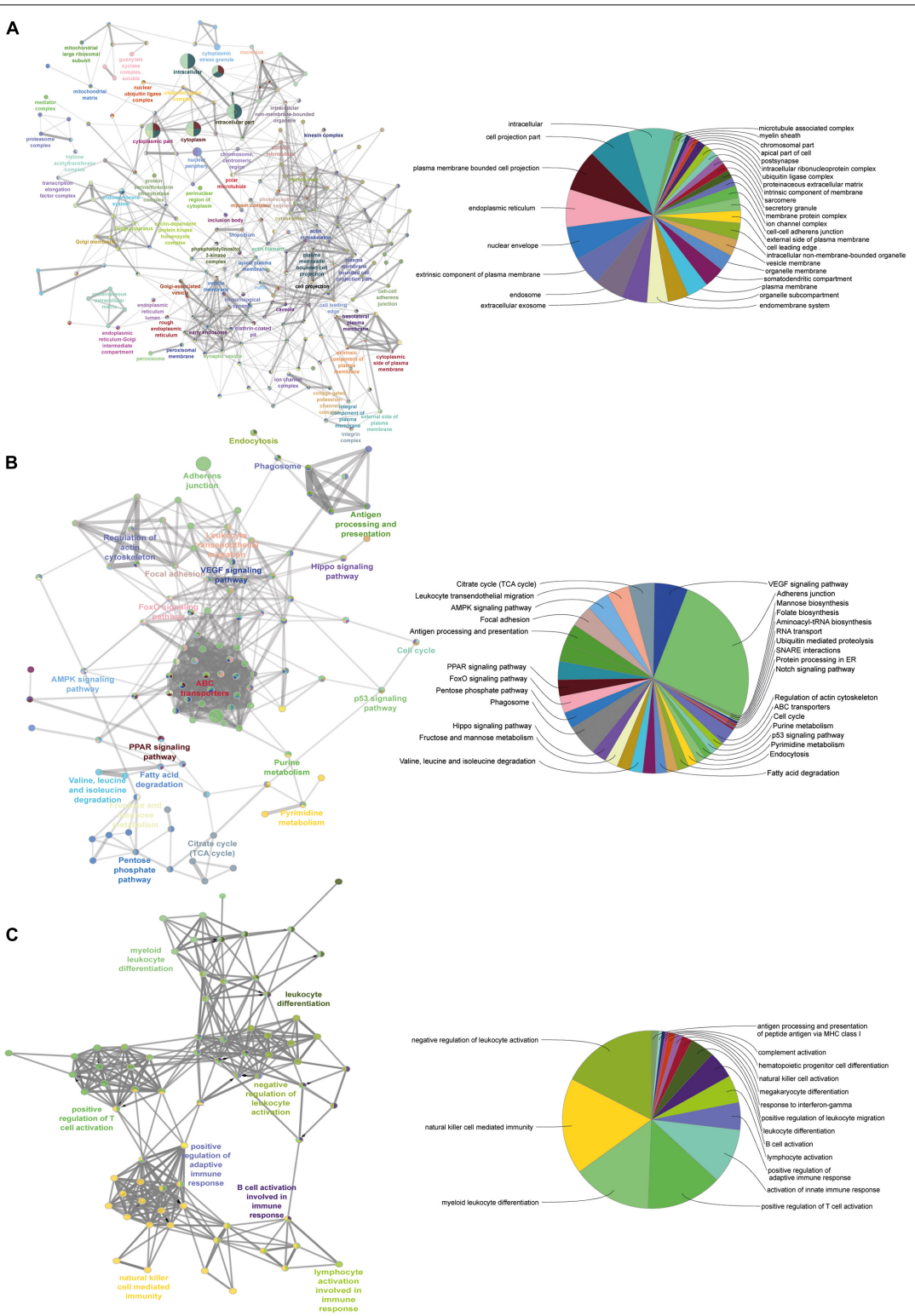


FIGURE 6 | Pathway–pathway interaction analysis. A functionally grouped network of enriched categories was generated for the target genes. GO terms are represented as nodes, and the node size represents the term enrichment significance. Functionally related groups partially overlap. The node pie charts represent the molecular function, immune system processes, reactome analysis of targets. Only the most important term in the group was labeled. **(A)** Representative molecular function interactions among targets. **(B)** Representative reactome analysis interactions among predicted targets. **(C)** Representative immune system processes interactions among targets.

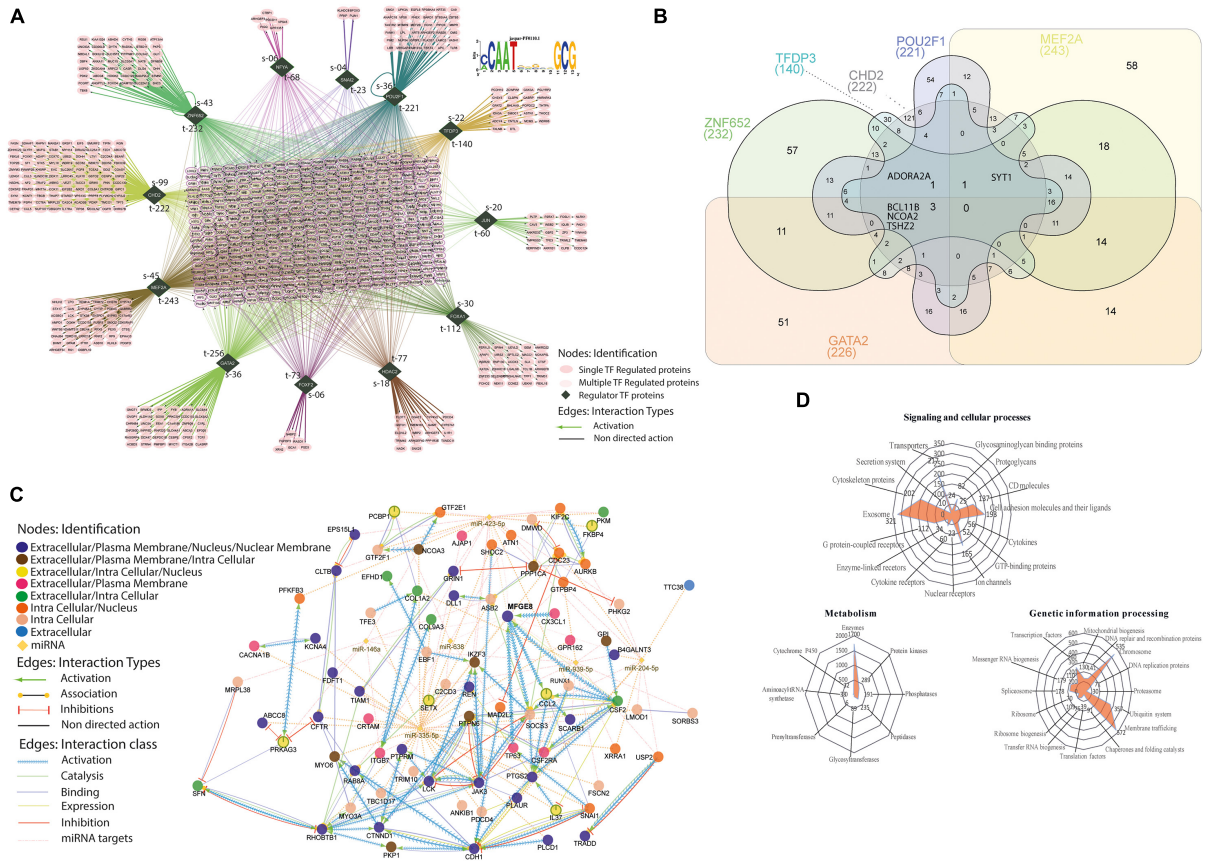


FIGURE 7 | Transcription factor(s) associated with regulome. **(A)** All the identified significantly ($p < 0.001$) differentially expressed proteins with 1.5-fold change up and down were used, $n = 938$ proteins for the identification of associated TFs involved in the cell cycle regulation. The dark green rhombus shape nodes show the TFs, *S* defines the number of singly associated protein partners, *t* defines the total number of protein partners associated with individual TFs. Interactions are color-coded and detailed described in the bottom right downside. JASPAR based position-specific scoring matrices predicted the regulatory control sequences that represent the binding specificity of TFs. **(B)** Venn Diagram. For the comparison of six TFs and their identified number of target proteins. It determined the common targets proteins associated with six TFs (GATA2, ZNF652, TFDP3, CHD2, POU2F1, and MEF2A), the total of targets in individual TF is reported below in round bracket. Inside the Venn diagram, the number reports the common proteins. **(C)** A directed network of MFGE8–miRNA interactome. The *in silico* based network analysis of directly interacted miRNAs and their associated proteins partners, which was also identified as significantly ($p < 0.001$) differentially expressed in MFGE8 KD proteome. See also **Supplementary Figure 5** for detailed analysis. **(D)** Radar plot. The identified proteins were mapped against the KEGG database for the determination of redundant involvement in molecular signaling and cellular processes, metabolic process, and genetic information processing. All the mapped proteins in the specific process were plotted in three different radar plots and indicated with the number of proteins determined.

for transcription regulatory region DNA binding (GO: 0044212, *p*-value-0.0022).

The affected proteins with these TFs further narrowed down to control cell cycle regulation and maintenance (GO: 0007050, cell cycle arrest, *p*-value-0.004; GO: 0051301, cell division, *p*-value-0.021). Secondly, we discover all the targets were identified to be in direct connection with the individual TF; instead, a large portion was synergistically regulated, and extensive overlap resulted in the formation of the large regulon in the center of the interactome (**Figure 7A**). Particularly, among the 226 targets proteins that were activated downstream to GATA2, we found a combination of multiple TFs controlled 77.5% of the total targets. However, interestingly only three proteins (BCL11B, NCOA2 TSHZ2) were associated with all the TFs, and 51 proteins were associated with the only GATA2, as shown in the logic diagram (**Figure 7A**). The same is the case with MEF2A (243 total, 76.2% commonly regulated, 58 unique targets), POU2F1 (221 total, 75.6% commonly regulated, 54 unique targets), CHD2 (222 total, 45.5% commonly regulated, 121 unique targets), TFDP3 (140 total, 78.6% commonly regulated, 30 unique targets), and ZNF652 (232 total, 75.5% commonly regulated, 57 unique targets) (see **Supplementary Table 7** for extensive details of the targets). These secondary regulons present the extensive overlap with the primary TFs regulon, intimating that these combinational functions of TFs are the leading contributors in gene regulation downstream to MFG-E8 (**Figure 7B**). The sharing of collective targets (more than 70% of each regulon) with the MFG-E8 regulon symbolizes the prevailing co-factorship among these TFs, something that has been described previously for individual TF, although not on such a lengthened order for the cell cycle proliferation. In extension, all of the TFs shown are implicated in cell cycle progression that is bona fide regulators

of cell growth progression (Kumar et al., 2006; Koga et al., 2007; Qiao et al., 2007; Semba et al., 2017; Tang et al., 2018; Wang et al., 2018).

Exploration of the Mechanism of MFG-E8 for Downstream Signaling

Next, we sought to determine the impact of MFG-E8 mediated signaling on cellular outcomes. For that, we also considered the possibility of a critical confounding variable of miRNA for the regulation and downstream signaling of MFG-E8 (**Figure 7B**). The combination of databases such as GeneMANIA, Reactome, CluePedia showed the activation of MFG-E8 through CX3CL1, TP63, and CSF2 in identified DEPs. The inferred directed network also indicates that MFG-E8 leads to activation of two proteins SOCS3 and CCL2, along with six miRNAs (miR-423-5p, miR-638, miR-939-5p, miR-204-5p, miR146a, and miR-335-5p) in the direct connection (**Supplementary Figures 7**, **8**). The miR-423-5p and miR-638 were identified for autophagy promotion, while miR-204p and miR 335-5p were previously reported in invasion and metastasis suppression. The interaction is expanded by the finding of associated proteins from DEPs with miRNAs and MFG-E8 collectively and divided it into four underlying strata that are extracellular, plasma membrane, intracellular, and nucleus associated. However, the network is described with a particular interaction type for all diverse associations such as activation, association, inhibition, catalysis, binding, and expression indicated with an individual variety of edges (**Figures 7B–D**). It helped us to identify the specific molecules regulating cellular growth via down-regulated MFG-E8 such as AURKB, TRADD, SNAI1, USP2, and KIF2C. **Figure 5** describes the illustrated signaling.

Determination of Cell Proliferation and Caspase Activity in MFG-E8 KD_MEC Cells

The downregulation of MFG-E8 protein resulted in the slow growth of the cells in MFG-E8 KD_MEC cells (**Figures 8A–C**). MTT assay results indicated that there was a decline in the MTT reductive capacity in MFG-E8 KD_MEC cells over different time points in culture (12, 24, 32, 48, 60, and 72 h) (**Figure 8A**), which was also validated by BrdU assay (**Figure 8B**). The cell growth rate was relatively equal for 12 and 24 h (non-significant $p > 0.05$). In comparison, it continuously declined from 32 to 72 h onward to nearly half with significant $p > 0.001$ (**Figure 8B**). Interestingly, these cellular proliferation assay showed a high growth rate of control cells in comparison with the MFG-E8 KD_MEC cells, especially at 72 h of cell incubation. To check whether MFG-E8 knockdown is leading to apoptosis in MEC cells, we performed the caspase 3/7 activity assay for all the same 6-time intervals (**Figure 8C**). Caspase 3/7 activity results were identified as non-significant changes for nearly all time points, suggesting that the decrease in the proliferation shown in cell growth is because of knockdown of MFG-E8 and not due to the induction of apoptosis. The findings were also confirmed through the presence of 11 different anti-apoptotic class proteins or lower expression of caspase proteins fold change was perceived (**Supplementary Table 9**). The results point out that MFG-E8 knockdown does not lead to caspases mediated apoptotic pathway. Similarly, MFG-E8 downregulation remarkably repressed wound healing and cell migration which is also highlighted in the enrichment analysis GO: 0016477, GO: 0043536: cell migration (p-value 0.000104) were validated by the migration assay ($P <$ value 0.0001) WT versus WT + HA14-1/KD/KD + HA14-1 proves our finding of proteomics-based pathway analysis data.

Dynamics of proteins phosphorylation majorly regulate the cellular signaling. From the network analysis, we selected three key TFs (AKT1, STAT3, and STAT5) and counterpart phosphoproteins for western blot analysis. We identified the non-significant changes in the non-phosphorylated AKT1, STAT3 and STAT5 proteins. However, significant upregulation of p-AKT (p-S473) and p-STAT3 (pS727) in KD cells were observed, but surprisingly, p-STAT5 (p-T694) was decreased in KD cells in comparison to WT cells (**Figure 8D**). The results are in correspondence to the proteomics data. Collectively, these data support that the dynamic phosphorylation event is more critical for the regulation of downstream signaling (**Figure 8E**).

DISCUSSION

The current study determined the signaling of MFG-E8 in MECs and its involvement in cellular proliferation and homeostasis. MFG-E8 protein regulates apoptotic-based cell death, especially the clearance of dying epithelial cells (Clarkson et al., 2004; Stein et al., 2004). The accumulation of milk at the end of the lactation stage in mammary glands triggers the death of excessive epithelial cells. These cells are promptly removed by neighboring epithelial cells and by infiltrating phagocytes with the help of MFG-E8 protein to ensure that the gland returns to its pre-pregnant state (Ensslin and Shur, 2007; Monks et al., 2008; Watson and Kreuzaler, 2011; Sumbal et al., 2020). Therefore, we focused on determining the pathways associated with MFG-E8 protein in MEC proliferation and oncogenesis processes.

Initially, our global sequence analysis across the species showed the involvement of MFGE8 protein in vesicle trafficking and membrane-associated ion channels, which functions for intracellular transportation (Peng and Elkon, 2011). Further, we found that the downregulation of MFG-E8 *in vitro* resulted in impaired epithelial morphology and decreased cell growth. The reduced cell growth rate is confirmed by cell growth confluence assay, wound healing and migration assay. The reason could be the lower expression of MFG-E8 is associated with a high level of cyclin-dependent kinase 5 activator protein (CDK5R2, FC-2), resulting in reduced cell growth. A similar observation was seen where MFG-E8 silencing in 4T1 cells reduces the cyclin protein expression level (Carrascosa et al., 2012). Nonetheless, MFG-E8 overexpression in MEC showed aggressive tumor growth *in vivo* (34).

Our cell cycle analysis in WT versus KD cells showed the increased cell cycle progression in KD cells. The reason could be the differential expression of cyclin D1, which regulates cell cycle progression during the G1/S transition. Its expression is increased

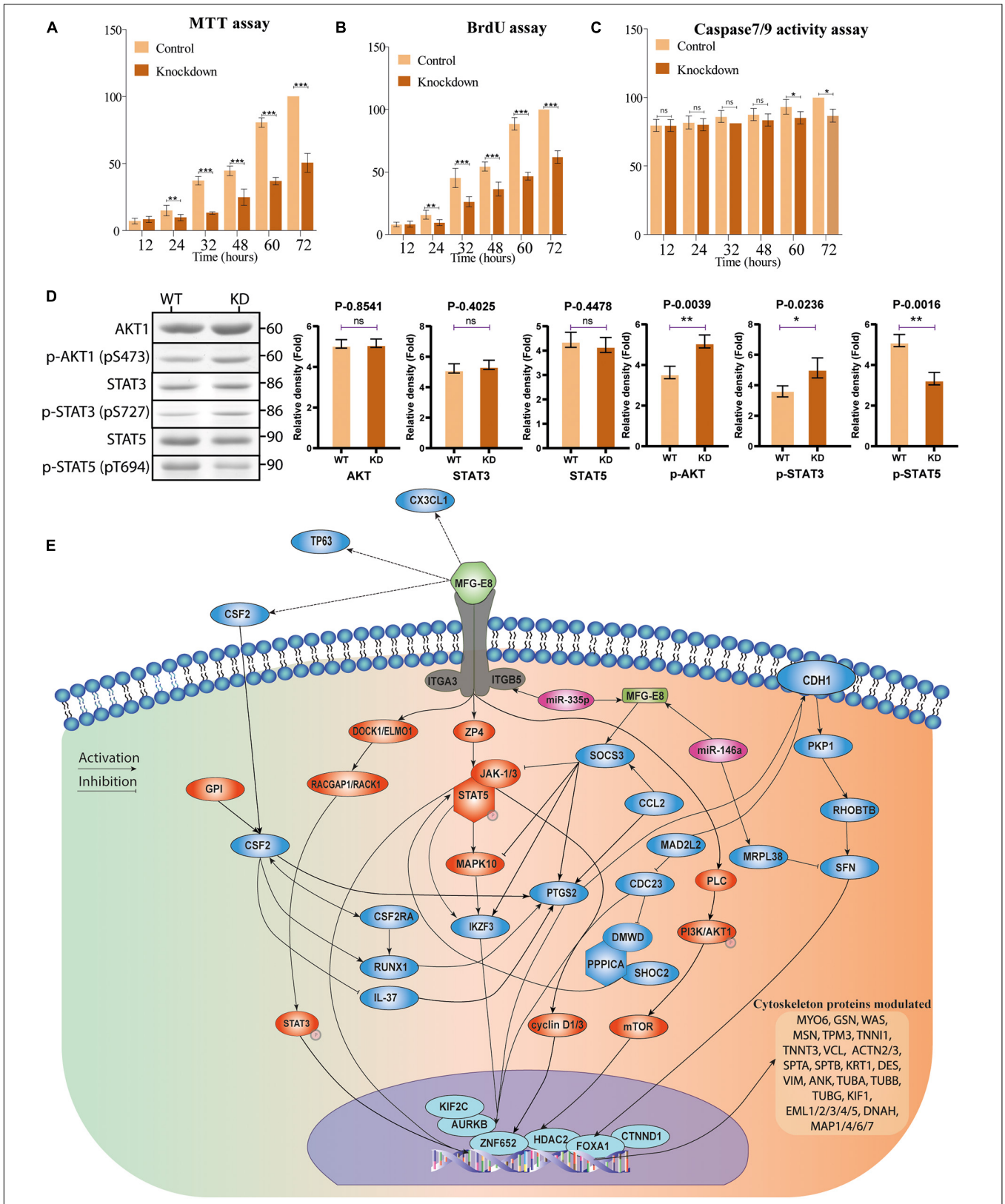


FIGURE 8 | Multiple assays for validation. **(A)** MTT assay. The cell proliferation MTT assay for the control and MFG-E8 KD BuMEC cells were performed. All experiments were performed six times points at 12-h intervals up to 72 h of culture. The values were calculated using a non-parametric *t*-test, and the error bar is

(Continued)

FIGURE 8 | Continued
shown with a standard error of the mean (SEM). $^{**}p < 0.01$, $^{***}p < 0.001$. **(B)** BrdU assay. The cell proliferation BrdU assay was assessed for the proliferation of the control and MFG-E8 KD BuMEC cells on every 12-h interval up to 72 h of culture by BrdU labeling. The values were calculated using a non-parametric *t*-test, and the error bar is shown with the standard error of the mean (SEM). $^{**}p < 0.01$, $^{***}p < 0.001$. **(C)** Caspase-3/7 assay. Caspase-3/7 activities of the control and MFG-E8 KD BuMEC cells were measured by Caspase-Glo 3/7 activity assay kit. The cells were seeded into 96-well plates and incubated for 12 to 72 h with an interval of 12 h each. The Caspase-Glo 3/7 reagent was added and incubated for another 3 h. Fluorescence intensities of the wells were then measured using excitation wavelength 492 nm and emission wavelength 535 nm. The caspase activities of the normal WT BuMEC cells were compared to transfected KD BuMEC cells. All experiments were performed six times with almost similar results, and values were calculated using a non-parametric *t*-test, and the error bar is shown with the standard error of the mean (SEM). $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.001$ and ns: non-significant. **(D)** Western blot analysis of scrambled shRNA WT and KD cells for AKT, STAT3 and STAT5 proteins with respective phosphoprotein antibody counterpart. Bar graph represent ± SD from three replicate. $^{*}p < 0.05$, $^{**}p < 0.005$, $^{***}p < 0.001$ and ns, non-significant. **(E)** Schematic representation for the signaling regulated through MFGE8 protein in mammary epithelial cells.

at the activation of oncogenes such as Ras or Neu and found to be up-regulated in breast cancer (Lee and Yang, 2003; Roy and Thompson, 2006). The transgenic mice overexpressing the cyclin D1 protein in the mammary gland resulted in hyperplasia and further led to neoplastic transformation (Wang et al., 1994). However, knockout in transgenic mice producing oncogenic protein Ras prevents mammary tumor development (Yu et al., 2001). Down-regulation of MFG-E8 resulting in low expression of cyclin D1 in MEC suggests that it is a valuable player involved in the initiation of MEC transformation, pushing it toward slowing down the cell growth; that is an essential requirement during the involution process.

Furthermore, in our iTRAQ based quantitative proteomic dataset, the network analysis reveals direct molecular linkage between knocked down state MFG-E8 activating ZP4 and STAT3/5 proteins. The STAT family proteins are orchestrators of mammary gland remodeling via cell proliferation and involution (Resemann et al., 2014). These TFs are present in the downstream signaling through various receptors, including cytokines receptors, growth factor receptors, and G-protein-coupled receptors. Here, we found that the ZP4 is physically interacting with JAK1/3-STAT5 and ITGA3/ITGB5 complex. Previously, the immunoprecipitation based results validate our findings, where zona pellucida-like domain-containing protein 1 (CUZD1), a similar protein like ZP4 in MECs, was found physically associated JAK-STAT5 complex (Mapes et al., 2017). The amplification of the CUZD1/ZP4 expression is associated with increased JAK1/JAK2/STAT5 signaling, which translocates the STAT5 into the nucleus, but the mechanism is poorly understood (Mapes et al., 2017). We determined that ZP4 potentiates JAK/STAT signaling downstream of MFG-E8 by mediating ITGA3/ITGB5 receptor activation by our bioinformatics data bridging an adaptor that assists in the recruitment of STAT5 to the stabilizing JAK1/3-STAT5 complex. The available literature about the roles of the effector proteins is the primary precedence of this hypothesis that alters signaling through this complex (Rawlings et al., 2004; Seif et al., 2017). An interesting example is a c-Src protein, which has been shown to propagate PRL initiated JAK/STAT signaling in healthy mammary tissue (García-Martínez et al., 2010). Likewise, STAT5 signaling is inhibited by the caveolin-1 (Cav-1) pathway by competitively binding to the domain of JAK2 tyrosine kinase, preventing subsequent activation of STAT5 (Park et al., 2002).

Another molecule regulated by MFG-E8 downstream signaling is DOCK1 resulted in the activation of RACGAP1 (FC-0.8)/RACK1 (FC-0.54) complex by STAT3 phosphorylation on Y705 (Tonozuka et al., 2004) and further transfer to the nucleus to act as a nuclear chaperone (Kawashima et al., 2009). The activation of Rac1 signaling supports the expression level of IL-6 production (Arulanandam et al., 2009), which is complementary to our results were down-regulation of this pathway results in the decreased expression of interleukins (with respective FCs; IL-6, IL-9, IL-16 0.52, IL-18 0.47, and IL-21). Previously, it was observed that the absence of DOCK1 in small mammary gland neoplasia reduced the STAT3 activation (Laurin et al., 2013). Here, we showed that the expression of STAT3 is diminished in the absence of MFG-E8 along with RAC1 and DOCK1. These results suggest that DOCK1 is a key GEF activating Rac1 complex and initiates the STAT3 activation mutually in mammary gland involution and breast cancer. It will be interesting to investigate the molecular intermediates of the RACGAP1/RACK1 complex *in vivo* to connect to STAT3 signaling pathway mammary gland developmental physiology as the down-regulation of this signaling resulted in the delays of the proliferation of MEC growth (**Figure 8E**).

On the other hand, we revealed that MFG-E8 mediated DOCK1/Rac1 signaling is essential for non-professional epithelial phagocytes to endorse the engulfment of apoptotic cells. This work is the first demonstration to address possible global conservation of the function of MFG-E8 mediated DOCK1/RAC1 signaling in dead cells clearance. The DOCK1-binding partner ELMO1 was formerly revealed to operate downstream to the PtdSer receptor BAIAP3 (a form of Bai1) to support the removal of dead cells. An *in vivo* study of ELMO1 mutant had shown a defect in removing apoptotic germ cells in the testis (Park et al., 2007; Elliott et al., 2010). Based on these results, we further confirm that BAIAP3 is necessarily required by epithelial cells to promote activation of the MFG-E8/DOCK1/ELMO1/RAC1 component to stimulate the elimination of dead MECs with respect to mammary gland involution. Simultaneously, the other phagocytic receptors, for example, ITGA3/ITGB5 or Tyro3, Axl, or MerTK, are equally vital to clear dying cells through DOCK1/RAC1 signaling. We reason that MFG-E8 mediated signaling is responsible for the activation of the DOCK1/RAC pathway, mainly required for the clearance of dying MECs.

The importance of the PI3K/AKT pathway is typically indispensable, and it plays a crucial role in the process of cell proliferation, survival, motility, secretion, apoptosis, and metabolism (Freudlsperger et al., 2011; Kamal et al., 2016).

The activation of the AKT pathway supports cell proliferation, but only limited reports have clearly explained the relationship between MFG-E8 and PI3K/Akt signaling (Neutzner et al., 2007; Jinushi et al., 2008). Our results indicate that downregulation of MFG-E8 leads to decreased expression of multiple PLC proteoforms mediated PI3K/AKT (FC 0.70) signaling via mTOR modulating the expression of TFs which resulted in the slow proliferation of the MEC cells. The results from the MTT, BrdU, and cell growth confluence assay suggested that the down-regulation of MFG-E8 reduces the cell proliferation rate and does not induce apoptosis, as confirmed by the caspase 3/7 activity assay. Our results of shRNA-based knockdown of the MFG-E8 in MECs cells are an excellent reference for future clinical trials. It was previously shown that preventing MFG-E8 signaling via anti- MFG-E8 antibodies caused the regression of experimental breast cancers (Ceriani et al., 1987). Concomitantly, using this approach, synergistically and traditional chemotherapy also showed reduced experimental colon carcinomas, melanomas, and lymphomas (Jinushi et al., 2009). MFG-E8 silenced triple-negative breast cancer cells showed higher sensitivity to cisplatin treatment (Yang et al., 2011). In this regards, collectively present results and published evidence, promoting the notion that MFG-E8 is a constitutive molecule essential for cellular homeostasis. Its expression pattern changes during mammary gland development to restore the molecular balance. This study also highlights that combinatorial therapies combined with the blocking of MFG-E8 may prove sufficiently effective outcomes in breast cancer treatment.

CONCLUSION

In this study, we utilized the benefits of a global proteomic workflow for measuring the crucial low abundant endogenous TFs network of signaling involved in cellular homeostasis downstream to MFG-E8 knockdown. Specifically, we observed that the absence of MFG-E8 suppresses cell growth, alters cellular morphology, and activates the immune function. Necessarily, suppression of MFG-E8 results in reduced growth of MECs by an autocrine/paracrine mechanism that may imply the induction of TFs such as cyclins D1, FOXA1, ZNF652, AURKB and many others for modulating the cytoskeleton-associated proteins and cell cycle regulators. The significance of MFG-E8 signaling concerning mammary gland developmental remodeling and disease is indisputable, and this study, through proteomic examination, has underscored its importance in cellular homeostasis. Finally, we anticipate that our comprehensive datasets will serve as a valuable foundation for further validation and mechanistic exposition of many novel proteins that are involved in MFG-E8 meditated signaling in the mammary gland.

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

SA and PS designed and performed the experiments, conducted data analysis, and wrote the manuscript. AV helped in initializing the experiments. SK helped in the data analysis. AM supervised the whole project and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We are thankful to DST-SERB for providing funding project number EMR/2017/000152. SA is thankful to the Indian Council of Medical Research (ICMR) for the SRF fellowship. PS is thankful to the Department of Biotechnology (DBT) for the SRF fellowship. We would also like to thank the Director of the ICAR-National Dairy Research Institute (NDRI), India, for providing the necessary facilities to carry out research work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.568660/full#supplementary-material

Supplementary Figure 1 | Full-length multiple sequence alignment of all 35 organisms with %Identity wise color code. Total numbers of amino acid residue present in the individual proteins sequence in all organisms are reported with the name. The consensus sequence logo was shown in the downside of the alignment, and the conserved nature was depicted using the size of the amino acid single alphabet symbol. Conservation and quality are reported using histograms. The highest histogram defines high conservation, and small histograms are low conservation.

Supplementary Figure 2 | Statistically significantly enriched motifs identified using multiple sequence alignment (MSA) of MFGE8 proteins retracted from Motif-X analysis. Motif-X online tool was used to identify the motifs of six amino acid residues for the identification of two specific post-translational modification (PTM) motif sequences. The amino acid residues are STY for phosphorylation and SWRN for glycosylation. The parameters used for the Motif-X algorithm were occurrences, 10; significance level, 0.00001.

Supplementary Figure 3 | Multi scatters plot. The graph represents the correlation of all the samples to all for the identification of the variable dependency.

Supplementary Figure 4 | Correlation and density plot. The iTRAQ intensities obtained after the MS data extraction was plotted to find the correlation among the replicates and between the conditions. The Pearson correlation *R* values were determined and reported in the graph. The blue color defines the high density of the spectra for the proteins and while the green color shows low, dense proteins.

Supplementary Figure 5 | The relative abundance of various signaling pathways based on DEPs iTRAQ data using the KEGG database (shown as that of KD versus WT cells). Gene Ontology is represented as fold enrichment. Probability was determined using a binomial statistic for the false discovery rate (FDR) and a *p*-value cutoff of % 0.05 significance level.

Supplementary Figure 6 | FACS analysis. Analysis of the BuMEC cell population before and after a knockdown by induces cell cycle arrest. **(A)** Flow cytometry analysis of the MFG-E8 KD cell cycle reveals that knockdown induces cell cycle arrest at the G0-G1 phase: 32.62%, G2-M phase: 20.18%, and S phase:47.20%. **(B)** While the analysis of the MFG-E8 WT cell cycle reveals induces cell cycle arrest at the G0-G1 phase: 18.42%, G2-M phase: 16.10%, and S phase: 65.85%.

Supplementary Figure 7 | Protein miRNA network of MFGE8–miRNA interactome. The *in silico* based network analysis of directly interacted miRNAs and their associated proteins partners, which was also identified as significantly ($p < 0.001$) differentially expressed in MFGE8 KD proteome. Circular nodes denote the proteins, diamond-shaped nodes are regulation miRNA partners. Dotted lined denotes the edges for miRNA to protein interactions, whereas the broad green line denotes the miRNA targets.

Supplementary Table 1 | Total combined identified protein and peptide list together with the identification of all ontological information associated in the Uniprot database.

Supplementary Table 2 | Classification of BuMEC proteome using PANTHER (Protein ANalysis THrough Evolutionary Relationships) curated biological database for protein families and their functionally related subfamilies to classify and identify the function of proteins.

Supplementary Table 3 | The resulted list of DAVID (the database for annotation, visualization, and integrated discovery) Functional Annotation Tool and typical protein-term enrichment analysis.

Supplementary Table 4 | KEGG BRITE annotation of MFGE8 linked differentially expressed proteins.

Supplementary Table 5 | Information from the ngLOC prediction of proteins classified in to different categories of cellular organelles.

Supplementary Table 6 | Information from string enrichment analysis for all the associated proteins involved in network construction.

Supplementary Table 7 | The determined information for the targets of TFs identified associated proteins in MFGE8 downstream signaling.

Supplementary Table 8 | Information on miRNA target proteins and interaction analysis for all the associated proteins involved in network construction.

Supplementary Table 9 | The information for the identified DEPs associated with apoptosis activity in the proteomics dataset.

REFERENCES

Abrahams, V. M., Kim, Y. M., Straszewski, S. L., Romero, R., and Mor, G. (2004). Macrophages and apoptotic cell clearance during pregnancy. *Am. J. Reprod. Immunol.* 4, 275–282. doi: 10.1111/j.1600-0897.2004.00156.x

Ali, S. A., Kaur, G., Kaushik, J. K., Malakar, D., Mohanty, A. K., and Kumar, S. (2017). Examination of pathways involved in leukemia inhibitory factor (LIF)-induced cell growth arrest using label-free proteomics approach. *J. Proteomics* 168, 37–52. doi: 10.1016/j.jprot.2017.07.014

Ali, S. A., Malakar, D., Kaushik, J. K., Mohanty, A. K., and Kumar, S. (2018). Recombinant purified buffalo leukemia inhibitory factor plays an inhibitory role in cell growth. *PLoS One* 13:e0198523. doi: 10.1371/journal.pone.0198523

Arulanandam, R., Vultur, A., Cao, J., Carefoot, E., Elliott, B. E., Truesdell, P. F., et al. (2009). Cadherin-cadherin engagement promotes cell survival via Rac1/Cdc42 and signal transducer and activator of transcription-3. *Mol. Cancer Res.* 8, 1310–1327. doi: 10.1158/1541-7786.MCR-08-0469

Atabai, K., Fernandez, R., Huang, X., Ueki, I., Kline, A., Li, Y., et al. (2005). Mfge8 is critical for mammary gland remodeling during involution. *Mol. Biol. Cell* 12, 5528–5537. doi: 10.1091/mbc.e05-02-0128

Aziz, M., Yang, W. L., Corbo, L. M., Chaung, W. W., Matsuo, S., and Wang, P. (2015). MFG-E8 inhibits neutrophil migration through αvβ3-integrin-dependent MAP kinase activation. *Int. J. Mol. Med.* 1, 18–28. doi: 10.3892/ijmm.2015.2196

Birge, R. B., Boeltz, S., Kumar, S., Carlson, J., Wanderley, J., Calianese, D., et al. (2016). Phosphatidylserine is a global immunosuppressive signal in efferocytosis, infectious disease, and cancer. *Cell Death Differ.* 6, 962–978. doi: 10.1038/cdd.2016.11

Carmon, L., Bobilev-Priel, I., Brenner, B., Bobilev, D., Paz, A., Bar-Haim, E., et al. (2002). Characterization of novel breast carcinoma–associated BA46-derived peptides in HLA-A2. 1/D b-β2mtransgenic mice. *J. Clin. Invest.* 4, 453–462. doi: 10.1172/JCI200214071

Carrascosa, C., Obula, R. G., Missiaglia, E., Lehr, H. A., Delorenzi, M., Frattini, M., et al. (2012). MFG-E8/lactadherin regulates cyclins D1/D3 expression and enhances the tumorigenic potential of mammary epithelial cells. *Oncogene* 12, 1521–1532. doi: 10.1038/onc.2011.356

Ceriani, R. L., Blank, E. W., and Peterson, J. A. (1987). Experimental immunotherapy of human breast carcinomas implanted in nude mice with a mixture of monoclonal antibodies against human milk fat globule components. *Cancer Res.* 2, 532–540. doi: <doi>

Ceriani, R. L., Peterson, J. A., Lee, J. Y., Moncada, R., and Blank, E. W. (1983). Characterization of cell surface antigens of human mammary epithelial cells with monoclonal antibodies prepared against human milk fat globule. *Somatic Cell Genet.* 4, 415–427. doi: 10.1007/BF01543043

Chopra, A., Ali, S. A., Bathla, S., Rawat, P., Vohra, V., Kumar, S., et al. (2020). High-resolution mass spectrometer–based ultra-deep profile of milk whey proteome in Indian Zebu (Sahiwal) cattle. *Front. Nutr.* 7:150. doi: 10.3389/fnut.2020.00150

Clarkson, R. W., Wayland, M. T., Lee, J., Freeman, T., and Watson, C. J. (2004). Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast Cancer Res.* 2:R92. doi: 10.1186/bcr754

Deutsch, E. W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., et al. (2010). A guided tour of the trans-proteomic pipeline. *Proteomics* 6, 1150–1159. doi: 10.1002/pmic.200900375

Elliott, M. R., Zheng, S., Park, D., Woodson, R. I., Reardon, M. A., Juncadella, I. J., et al. (2010). Unexpected requirement for ELMO1 in clearance of apoptotic germ cells in vivo. *Nature* 467, 333–337. doi: 10.1038/nature09356

Ensslin, M. A., and Shur, B. D. (2007). The EGF repeat and discoidin domain protein, SED1/MFG-E8, is required for mammary gland branching morphogenesis. *Proc. Natl Acad. Sci. U.S.A.* 8, 2715–2720. doi: 10.1073/pnas.0610296104

Freudlsperger, C., Burnett, J. R., Friedman, J. A., Kannabiran, V. R., Chen, Z., and Van Waes, C. (2011). EGFR–PI3K–AKT–mTOR signaling in head and neck squamous cell carcinomas: attractive targets for molecular-oriented therapy. *Expert Opin. Ther. Targets* 1, 63–74. doi: 10.1517/14728222.2011.541440

García-Martínez, J. M., Calcabrini, A., González, L., Martín-Forero, E., Agulló-Ortuño, M. T., Simon, V., et al. (2010). A non-catalytic function of the Src family tyrosine kinases controls prolactin-induced Jak2 signaling. *Cell. Signal.* 3, 415–426. doi: 10.1016/j.cellsig.2009.10.013

Hanayama, R., and Nagata, S. (2005). Impaired involution of mammary glands in the absence of milk fat globule EGF factor 8. *Proc. Natl. Acad. Sci. U.S.A.* 46, 16886–16891. doi: 10.1073/pnas.0508599102

Hanayama, R., Tanaka, M., Miwa, K., Shinohara, A., Iwamatsu, A., and Nagata, S. (2002). Identification of a factor that links apoptotic cells to phagocytes. *Nature* 6885, 182–187. doi: 10.1038/417182a

Hanayama, R., Tanaka, M., Miyasaka, K., Aozasa, K., Koike, M., Uchiyama, Y., et al. (2004). Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice. *Science* 5674, 1147–1150. doi: 10.1126/science.1094359

Hendricks, L., Aziz, M., Yang, W. L., Nicastro, J., Coppa, G. F., Symons, M., et al. (2017). Milk fat globule-epidermal growth factor-factor VIII–derived peptide MSP68 is a cytoskeletal immunomodulator of neutrophils that inhibits Rac1. *J. Surg. Res.* 208, 10–19. doi: 10.1016/j.jss.2016.08.098

Hidai, C., Zupancic, T., Penta, K., Mikhail, A., Kawana, M., Quertermous, E. E., et al. (1998). Cloning and characterization of developmental endothelial locus-1: an embryonic endothelial cell protein that binds the αvβ3 integrin receptor. *Genes Dev.* 1, 21–33. doi: 10.1101/gad.12.1.21

Hirano, Y., Yang, W. L., Aziz, M., Zhang, F., Sherry, B., and Wang, P. (2017). MFG-E8-derived peptide attenuates adhesion and migration of immune cells to endothelial cells. *J. Leukocyte Biol.* 5, 1201–1209. doi: 10.1189/jlb.3A0416-184RR

Hu, C. Y., Wu, C. S., Tsai, H. F., Chang, S. K., Tsai, W. I., and Hsu, P. N. (2009). Genetic polymorphism in milk fat globule-EGF factor 8 (MFG-E8) is associated with systemic lupus erythematosus in human. *Lupus* 8, 676–681. doi: 10.1177/0961203309103027

Janjanam, J., Singh, S., Jena, M. K., Varshney, N., Kola, S., Kumar, S., et al. (2014). Comparative 2D-DIGE proteomic analysis of bovine mammary epithelial cells during lactation reveals protein signatures for lactation persistency and milk yield. *PLoS One* 9:e102515. doi: 10.1371/journal.pone.0102515

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput. Biol.* 7:e1003731. doi: 10.1371/journal.pcbi.1003731

Jinushi, M., Draganov, D., Carrascco, D., Mihm, M., Tahara, H., and Dranoff, G. (2008). MFG-E8 derived from tumor-infiltrating myeloid cells plays a critical role in promoting tumor growth and metastasis. *Cancer Res.* 68:4198. doi: 10.1158/0008-5472.CAN-08-2147

Jinushi, M., Nakazaki, Y., Dougan, M., Carrasco, D. R., Mihm, M., and Dranoff, G. (2007). MFG-E8–mediated uptake of apoptotic cells by APCs links the pro-and antiinflammatory activities of GM-CSF. *J. Clin. Invest.* 7, 1902–1913. doi: 10.1172/JCI30966

Jinushi, M., Sato, M., Kanamoto, A., Itoh, A., Nagai, S., Koyasu, S., et al. (2009). Milk fat globule epidermal growth factor–8 blockade triggers tumor destruction through coordinated cell-autonomous and immune-mediated mechanisms. *J. Exp. Med.* 6, 1317–1326. doi: 10.1084/jem.20082614

Kamal, A., Nayak, V. L., Nagesh, N., Vishnuvardhan, M. V., and Reddy, N. S. (2016). Benzo [b] furan derivatives induces apoptosis by targeting the PI3K/Akt/mTOR signaling pathway in human breast cancer cells. *Bioorg. Chem.* 66, 124–131. doi: 10.1016/j.bioorg.2016.04.004

Kaur, G., Ali, S. A., Pachauri, S., Malakar, D., Kaushik, J. K., Mohanty, A. K., et al. (2017). Buffalo leukemia inhibitory factor induces differentiation and dome-like secondary structures in COS-1 cells. *Cytogenet. Genome Res.* 3, 119–130. doi: 10.1159/000465507

Kawashima, T., Bao, Y. C., Minoshima, Y., Nomura, Y., Hatori, T., Hori, T., et al. (2009). A Rac GTPase-activating protein, MgcRacGAP, is a nuclear localizing signal-containing nuclear chaperone in the activation of STAT transcription factors. *Mol. Cell. Biol.* 7, 1796–1813. doi: 10.1128/MCB.01423-08

Ko, D. S., Kim, S. H., Park, J. Y., Lee, G., Kim, H. J., Kim, G., et al. (2020). Milk fat globule-EGF factor 8 contributes to progression of hepatocellular carcinoma. *Cancers* 12:403. doi: 10.3390/cancers12020403

Koga, S., Yamaguchi, N., Abe, T., Minegishi, M., Tsuchiya, S., Yamamoto, M., et al. (2007). Cell-cycle–dependent oscillation of GATA2 expression in hematopoietic cells. *Blood* 10, 4200–4208. doi: 10.1182/blood-2006-08-044149

Kumar, R., Manning, J., Spendlove, H. E., Kremmidiotis, G., McKirdy, R., Lee, J., et al. (2006). ZNF652, a novel zinc finger protein, interacts with the putative breast tumor suppressor CBFA2T3 to repress transcription. *Mol. Cancer Res.* 9, 655–665. doi: 10.1158/1541-7786.MCR-05-0249

Laurin, M., Huber, J., Pelletier, A., Houalla, T., Park, M., Fukui, Y., et al. (2013). Rac-specific guanine nucleotide exchange factor DOCK1 is a critical regulator of HER2-mediated breast cancer metastasis. *Proc. Natl. Acad. Sci. U.S.A.* 18, 7434–7439. doi: 10.1073/pnas.1213050110

Lee, M. H., and Yang, H. Y. (2003). Regulators of G1 cyclin–dependent kinases and cancers. *Cancer Metastasis Rev.* 4, 435–449. doi: 10.1023/A:1023785332315

Li, H., Zhang, T., Wang, K., Lu, M., Guo, Y., Zhang, Y., et al. (2019). MFGE8 protects against CCl4-induced liver injury by reducing apoptosis and promoting proliferation of hepatocytes. *J. Cell. Physiol.* 234, 16463–16474. doi: 10.1002/jcp.28314

Lönnerdal, B. (2003). Nutritional and physiologic significance of human milk proteins. *Am. J. Clin. Nutr.* 77, 1537S–1543S. doi: 10.1093/ajcn/77.6.1537S

Lotfan, M., Ali, S. A., Yadav, M. L., Choudhary, S., Jena, M. K., Kumar, S., et al. (2018). Genome-wide gene expression analysis of 45 days pregnant fetal cotyledons vis-a-vis non-pregnant caruncles in buffalo (*Bubalus bubalis*). *Gene* 654, 127–137. doi: 10.1016/j.gene.2018.02.038

MacLean, B., Eng, J. K., Beavis, R. C., and McIntosh, M. (2006). General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* 22, 2830–2832. doi: 10.1093/bioinformatics/btl379

Mapes, J., Li, Q., Kannan, A., Anandan, L., Laws, M., Lydon, J. P., et al. (2017). CUZD1 is a critical mediator of the JAK/STAT5 signaling pathway that controls mammary gland development during pregnancy. *PLoS Genet.* 3:e1006654. doi: 10.1371/journal.pgen.1006654

Monks, J., Geske, F. J., Lehman, L., and Fadok, V. A. (2002). Do inflammatory cells participate in mammary gland involution? *J. Mammary Gland Biol. Neoplasia* 2, 163–176. doi: 10.1023/A:1020351919634

Monks, J., Rosner, D., Geske, F. J., Lehman, L., Hanson, L., Neville, M. C., et al. (2005). Epithelial cells as phagocytes: apoptotic epithelial cells are engulfed by mammary alveolar epithelial cells and repress inflammatory mediator release. *Cell Death Differ.* 2, 107–114. doi: 10.1038/sj.cdd.4401517

Monks, J., Smith-Steinhart, C., Kruk, E. R., Fadok, V. A., and Henson, P. M. (2008). Epithelial cells remove apoptotic epithelial cells during post-lactation involution of the mouse mammary gland. *Biol. Reprod.* 4, 586–594. doi: 10.1095/biolreprod.107.065045

Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 17, 4646–4658. doi: 10.1021/ac0341261

Neutzner, M., Lopez, T., Feng, X., Bergmann-Leitner, E. S., Leitner, W. W., and Udey, M. C. (2007). MFG-E8/lactadherin promotes tumor growth in an angiogenesis-dependent transgenic mouse model of multistage carcinogenesis. *Cancer Res.* 14, 6777–6785. doi: 10.1158/0008-5472.CAN-07-0165

Oshima, K., Aoki, N., Kato, T., Kitajima, K., and Matsuda, T. (2002). Secretion of a peripheral membrane protein. MFG-E8, as a complex with membrane vesicles: a possible role in membrane secretion. *Eur. J. Biochem.* 4, 1209–1218. doi: 10.1046/j.1432-1033.2002.02758.x

Oshima, K., Aoki, N., Negi, M., Kishi, M., Kitajima, K., and Matsuda, T. (1999). Lactation-dependent expression of an mRNA splice variant with an exon for a multiplyo-glycosylated domain of mouse milk fat globule glycoprotein MFG-E8. *Biochem. Biophys. Res. Commun.* 3, 522–528. doi: 10.1006/bbrc.1998.0107

Park, D. S., Lee, H., Frank, P. G., Razani, B., Nguyen, A. V., Parlow, A. F., et al. (2002). Caveolin-1-deficient mice show accelerated mammary gland development during pregnancy, premature lactation, and hyperactivation of the Jak-2/STAT5a signaling cascade. *Mol. Biol. Cell* 10, 3416–3430. doi: 10.1091/mbc.02-05-0071

Park, D., Tosello-Trampont, A. C., Elliott, M. R., Lu, M., Haney, L. B., Ma, Z., et al. (2007). BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. *Nature* 7168, 430–434. doi: 10.1038/nature06329

Peng, Y., and Elkon, K. B. (2011). Autoimmunity in MFG-E8–deficient mice is associated with altered trafficking and enhanced cross-presentation of apoptotic cell antigens. *J. Clin. Invest.* 6, 2221–2241. doi: 10.1172/JCI43254

Peterman, E., Gibieža, P., Schafer, J., Skeberdis, V. A., Kaupinis, A., Valius, M., et al. (2019). The post-abscission midbody is an intracellular signaling organelle that regulates cell proliferation. *Nat. Commun.* 10, 1–6. doi: 10.1038/s41467-019-10871-0

Qiao, H., Di Stefano, L., Tian, C., Li, Y. Y., Yin, Y. H., Qian, X. P., et al. (2007). Human TFDP3, a novel DP protein, inhibits DNA binding and transactivation by E2F. *J. Biol. Chem.* 1, 454–466. doi: 10.1074/jbc.M606169200

Rawlings, J. S., Rosler, K. M., and Harrison, D. A. (2004). The JAK/STAT signaling pathway. *J. Cell Sci.* 8, 1281–1283. doi: 10.1242/jcs.00963

Raymond, A., Ensslin, M. A., and Shur, B. D. (2009). SED1/MFG-E8: a bi-motif protein that orchestrates diverse cellular interactions. *J. Cell. Biochem.* 6, 957–966. doi: 10.1002/jcb.22076

Resemann, H. K., Watson, C. J., and Lloyd-Lewis, B. (2014). The Stat3 paradox: a killer and an oncogene. *Mol. Cell. Endocrinol.* 1, 603–611. doi: 10.1016/j.mce.2013.06.029

Roy, P. G., and Thompson, A. M. (2006). Cyclin D1 and breast cancer. *Breast* 6, 718–727. doi: 10.1016/j.breast.2006.02.005

Seif, F., Khoshmirsafa, M., Aazami, H., Mohsenzadegan, M., Sedighi, G., and Bahar, M. (2017). The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Commun. Signal.* 1:23. doi: 10.1186/s12964-017-0177-y

Semba, Y., Harada, A., Maehara, K., Oki, S., Meno, C., Ueda, J., et al. (2017). Chd2 regulates chromatin for proper gene expression toward differentiation in mouse embryonic stem cells. *Nucleic Acids Res.* 15, 8758–8772. doi: 10.1093/nar/gkx475

Silvestre, J. S., Théry, C., Hamard, G., Boddaert, J., Aguilar, B., Delcayre, A., et al. (2005). Lactadherin promotes VEGF-dependent neovascularization. *Nat. Med.* 5, 499–506. doi: 10.1038/nm1233

Stein, T., Morris, J. S., Davies, C. R., Weber-Hall, S. J., Duffy, M. A., Heath, V. J., et al. (2004). Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP. CD14 and STAT3. *Breast Cancer Res.* 2:R75. doi: 10.1186/bcr753

Stein, T., Salomonis, N., and Gusterson, B. A. (2007). Mammary gland involution as a multi-step process. *J. Mammary Gland Biol. Neoplasia* 1, 25–35. doi: 10.1007/s10911-007-9035-7

Stubbs, J. D., Lekutis, C., Singer, K. L., Bui, A., Yuzuki, D., Srinivasan, U., et al. (1990). cDNA cloning of a mouse mammary epithelial cell surface protein reveals the existence of epidermal growth factor-like domains linked to factor VIII-like sequences. *Proc. Natl. Acad. Sci. U.S.A.* 21, 8417–8421. doi: 10.1073/pnas.87.21.8417

Sugano, G., Bernard-Pierrot, I., Lae, M., Battail, C., Allory, Y., Stransky, N., et al. (2011). Milk fat globule—epidermal growth factor—factor VIII (MFGE8)/lactadherin promotes bladder tumor development. *Oncogene* 6, 642–653. doi: 10.1038/onc.2010.446

Suhail, A., Rizvi, Z. A., Mujagond, P., Ali, S. A., Gaur, P., Singh, M., et al. (2019). DeSUMOylase SENP7-mediated epithelial signaling triggers intestinal inflammation via expansion of gamma-delta T cells. *Cell Rep.* 29, 3522–3538. doi: 10.1016/j.celrep.2019.11.028

Sumbal, J., Chiche, A., Charifou, E., Koledova, Z., and Li, H. (2020). Primary mammary organoid model of lactation and involution. *Front. Cell Dev. Biol.* 8:68. doi: 10.3389/fcell.2020.00068

Tang, X., Zhao, Y., Buchon, N., and Engström, Y. (2018). The POU/Oct transcription factor nubbin controls the balance of intestinal stem cell maintenance and differentiation by isoform-specific regulation. *Stem Cell Rep.* 5, 1565–1578. doi: 10.1016/j.stemcr.2018.03.014

Taylor, M. R., Couto, J. R., Scallan, C. D., Ceriani, R. L., and Peterson, J. A. (1997). Lactadherin (formerly BA46), a membrane-associated glycoprotein expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp (RGD)-dependent cell adhesion. *DNA Cell Biol.* 7, 861–869. doi: 10.1089/dna.1997.16.861

Tonozuka, Y., Minoshima, Y., Bao, Y. C., Moon, Y., Tsubono, Y., Hatori, T., et al. (2004). A GTPase-activating protein binds STAT3 and is required for IL-6–induced STAT3 activation and for differentiation of a leukemic cell line. *Blood* 12, 3550–3557. doi: 10.1182/blood-2004-03-1066

Wang, T. C., Cardiff, R. D., Zukerberg, L., Lees, E., Arnold, A., and Schmidt, E. V. (1994). Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. *Nature* 369, 669–671. doi: 10.1038/369669a0

Wang, Y. N., Yang, W. C., Li, P. W., Wang, H. B., Zhang, Y. Y., and Zan, L. S. (2018). Myocyte enhancer factor 2A promotes proliferation and its inhibition attenuates myogenic differentiation via myozenin 2 in bovine skeletal muscle myoblast. *PLoS One* 13:e0196255. doi: 10.1371/journal.pone.0196255

Watson, C. J. (2006a). Key stages in mammary gland development-Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. *Breast Cancer Res.* 2:203. doi: 10.1186/bcr1401

Watson, C. J. (2006b). Post-lactational mammary gland regression: molecular basis and implications for breast cancer. *Expert Rev. Mol. Med.* 32, 1–5. doi: 10.1017/S1462399406000196

Watson, C. J., and Kreuzaler, P. A. (2011). Remodeling mechanisms of the mammary gland during involution. *Int. J. Dev. Biol.* 55, 757–762. doi: 10.1387/ijdb.113414cw

Wu, J., Yang, H., Cheng, J., Zhang, L., Ke, Y., Zhu, Y., et al. (2020). Knockdown of milk-fat globule EGF factor-8 suppresses glioma progression in GL261 glioma cells by repressing microglial M2 polarization. *J. Cell. Physiol.* 235, 8679–8690. doi: 10.1002/jcp.29712

Yamazaki, M., Maruyama, S., Abé, T., Tsuneki, M., Kato, H., Izumi, K., et al. (2020). Rac1-dependent phagocytosis of apoptotic cells by oral squamous cell carcinoma cells: a possible driving force for tumor progression. *Exp. Cell Res.* 392:112013. doi: 10.1016/j.yexcr.2020.112013

Yang, C., Hayashida, T., Forster, N., Li, C., Shen, D., Maheswaran, S., et al. (2011). The integrin αvβ3-5 ligand MFG-E8 is a p63/p73 target gene in triple-negative breast cancers but exhibits suppressive functions in ER+ and erbB2+ breast cancers. *Cancer Res.* 3, 937–945. doi: 10.1158/0008-5472.CAN-10-1471

Yu, Q., Geng, Y., and Sicinski, P. (2001). Specific protection against breast cancers by cyclin D1 ablation. *Nature* 6841, 1017–1021. doi: 10.1038/35082500

Zhao, Q. J., Yu, Y. B., Zuo, X. L., Dong, Y. Y., and Li, Y. Q. (2012). Milk fat globule-epidermal growth factor 8 is decreased in intestinal epithelium of ulcerative colitis patients and thereby causes increased apoptosis and impaired wound healing. *Mol. Med.* 3, 497–506. doi: 10.2119/molmed.2011.00369

Zhou, Y., Bond, A. M., Shade, J. E., Zhu, Y., Chung-ha, O. D., Wang, X., et al. (2018). Autocrine Mfge8 signaling prevents developmental exhaustion of the adult neural stem cell pool. *Cell Stem Cell* 23, 444–452. doi: 10.1016/j.stem.2018.08.005

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 Verma, Ali, Singh, Kumar and Mohanty. 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 Oncology

ORIGINAL RESEARCH
published: 14 May 2021
doi: 10.3389/fonc.2021.548243

Multiple Reaction Monitoring-Based Targeted Assays for the Validation of Protein Biomarkers in Brain Tumors

Saicharan Ghantasala[1†], Medha Gayathri J. Pai[2†], Deeptarup Biswas[2†], Nikita Gahoi[1], Shuvolina Mukherjee[2], Manubhai KP[2], Mehar Un Nissa[2], Alisha Srivastava[3], Sridhar Epari[4,5], Prakash Shetty[5,6], Aliasgar Moiyadi[5,6] and Sanjeeva Srivastava[2]*

[1] Centre for Research in Nanotechnology and Science, Indian Institute of Technology Bombay, Mumbai, India, [2] Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India, [3] Department of Genetics, University of Delhi, New Delhi, India, [4] Department of Pathology, Tata Memorial Centre's – Advanced Centre for Treatment, Research and Education in Cancer, Navi Mumbai, India, [5] Homi Bhabha National Institute, Mumbai, India, [6] Department of Neurosurgery, Tata Memorial Centre's – Advanced Centre for Treatment, Research and Education in Cancer, Navi Mumbai, India

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular & Molecular Biology (CCMB), India

Reviewed by:
Wendy Heywood, University College London, United Kingdom
Fabio Luis Forti, University of São Paulo, Brazil

***Correspondence:**
Sanjeeva Srivastava
sanjeeva@iitb.ac.in

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *01 April 2020*
Accepted: *19 April 2021*
Published: *14 May 2021*

Citation:
Ghantasala S, Pai MGJ, Biswas D, Gahoi N, Mukherjee S, KP M, Nissa MU, Srivastava A, Epari S, Shetty P, Moiyadi A and Srivastava S (2021) Multiple Reaction Monitoring-Based Targeted Assays for the Validation of Protein Biomarkers in Brain Tumors. Front. Oncol. 11:548243. doi: 10.3389/fonc.2021.548243

The emergence of omics technologies over the last decade has helped in advancement of research and our understanding of complex diseases like brain cancers. However, barring genomics, no other omics technology has been able to find utility in clinical settings. The recent advancements in mass spectrometry instrumentation have resulted in proteomics technologies becoming more sensitive and reliable. Targeted proteomics, a relatively new branch of mass spectrometry-based proteomics has shown immense potential in addressing the shortcomings of the standard molecular biology-based techniques like Western blotting and Immunohistochemistry. In this study we demonstrate the utility of Multiple reaction monitoring (MRM), a targeted proteomics approach, in quantifying peptides from proteins like Apolipoprotein A1 (APOA1), Apolipoprotein E (APOE), Prostaglandin H2 D-Isomerase (PTGDS), Vitronectin (VTN) and Complement C3 (C3) in cerebrospinal fluid (CSF) collected from Glioma and Meningioma patients. Additionally, we also report transitions for peptides from proteins – Vimentin (VIM), Cystatin-C (CST3) and Clusterin (CLU) in surgically resected Meningioma tissues; Annexin A1 (ANXA1), Superoxide dismutase (SOD2) and VIM in surgically resected Glioma tissues; and Microtubule associated protein-2 (MAP-2), Splicing factor 3B subunit 2 (SF3B2) and VIM in surgically resected Medulloblastoma tissues. To our knowledge, this is the first study reporting the use of MRM to validate proteins from three types of brain malignancies and two different bio-specimens. Future studies involving a large cohort of samples aimed at accurately detecting and quantifying peptides of proteins with roles in brain malignancies could potentially result in a panel of proteins showing ability to classify and grade tumors. Successful application of these techniques could ultimately offer alternative strategies with increased accuracy, sensitivity and lower turnaround time making them translatable to the clinics.

Keywords: targeted proteomics, gliomas, Meningioma, Medulloblastoma, multiple reaction monitoring

INTRODUCTION

Advances in Mass spectrometry have provided a major impetus to the field of proteomics over the last decade. These developments have resulted in proteomics playing a role in advancing our understanding of disease biology and opened new avenues such as biomarker development, augmentation of therapeutic modalities and drug discovery. Global consortia like CPTAC (Clinical Proteomics Tumor Analysis Consortium), TCGA (The Cancer Genome Atlas) and HUPO (Human Proteome Organization) have played a major role in utilising the power of omics technologies towards understanding the underlying mechanisms of various cancers (1–3). The development of new softwares, global repositories and reproducible workflows has also played a key role in increasing the utility of proteomics methodologies for research in the last decade. However, despite these advancements in technology and our understanding of cancer, early detection and treatment of malignancies related to the brain continue to pose a serious challenge even today (4). Primary brain tumors are a heterogeneous group of malignant and benign tumors arising from brain parenchyma or the cell types existing in the cranial vault including cranial nerves, meninges, neuroepithelial tissues, germ cells, pituitary gland, and even residual embryonic tissues. These tumors are characterized by high morbidity and mortality rates due to their localization, invasive growth and heterogeneous nature (5). Among all CNS tumors, glioma accounts for approximately 28% of tumors, of which 80% are malignant (6). Glioblastomas are the most prevalent and malignant sub-type of gliomas (3.20 per 100,000 population) followed by diffuse astrocytomas (0.48 per 100,000 population) (7). Meningiomas are slow-growing extra-axial lesions, originating from cells of arachnoid villi or meningothelial cells present in the meninges. These are most prevalent primary brain tumors and their incidences seem to rise with an increase in age. Medulloblastomas are embryonal tumors commonly seen in children. All medulloblastomas are classified pathologically as Grade IV due to their aggressive nature. Transcriptomic studies have identified four subtypes of medulloblastoma, which include SHH, WNT, G4 and G3, each with a distinct clinical and therapeutic presentation (8). The 2016 classification of WHO for brain cancers included molecular markers along with histological parameters for better clinical identification (9).

Recent advances in molecular biology and genomics have immensely benefitted the classification and prognosis of brain cancers. The omics research has also contributed in the brain tumour investigation which has unravelled different underlying biological mechanism, accelerated biomarker discovery (10–12). However, there still exits a gap between direct implementation of these studies in the clinics, as brain tumors are among the deadliest cancers. Successful translation of candidate biomarkers is also limited due to the unavailability of antibodies required for the validation experiments on large cohorts. The recognition of the potential of targeted proteomics approaches such as selected reaction monitoring, accelerated its use in hypothesis driven proteomics research (13). These proteomics approaches offer advantages such as high sensitivity and accuracy over the conventional validation approaches which rely on the use of antibodies, are labour intensive and time consuming. These approaches are slowly but surely finding use in biomedical research (14). Faria et al. has extensively reviewed the revolutionary role of targeted proteomics in cancer biomarker discovery and its shift from shotgun proteomics (15). Targeted proteomics has also showed its immense potential in the field of infectious disease, understanding metabolic disorders and in accelerating our understanding of the SARS CoV2 virus and its mode of infection during the ongoing pandemic (16–18). The integration of proteogenomics and targeted proteomic validation could be a promising tool owing to increased robustness, sensitivity and selectivity in cancer research (19). The Verification Working Group (VWG) of the CPTAC consortium carried out multiple experiment with an aim to assess the reproducibility, robustness, and transferability of MRM based assays. The findings from the study highlight the utility and practicality of highly reproducible MS-based assays and their potential role in clinics (20). In the current study, we have employed MRM as a proof of concept to validate a few proteins previously identified in our discovery studies on the three major brain malignancies - Gliomas, Meningiomas and Medulloblastomas. To our knowledge, this is the first study to validate potential biomarkers in brain tumors from CSF and tissue proteins using mass spectrometry-based MRM approach.

MATERIALS AND METHODS

Preparation of the Clinical Samples for Mass Spectrometric Analysis

Glioma and Meningioma CSF

Protein extraction from 17 CSF samples (MG I (n=3), MG II (n=4), GBM (n=5) and LGG or low-grade gliomas (n=5)) from Glioma and Meningioma patients was performed using the urea lysis protocol optimized in our lab. Please refer to **Supplementary Table 1** for clinical information about the samples used. 500µL of CSF from Meningioma and Glioma patients was concentrated to 200µL each and processed. The CSF samples were sonicated at an amplitude of 20% with 5 sec pulse for a total of 8 cycles. Following this, the samples were vacuum dried, and the pellet was reconstituted in 6M Urea buffer. Protein amount in the sample was then quantified using 2D quant kit (GE-Healthcare, Sweden). From this, 50µg of protein sample was reduced with 0.5M TCEP at a final concentration of 20mM for one hour at 37°C followed by alkylation using 37.5mM IAA for 30 minutes at RT. The samples were diluted with 25mM Tris buffer (pH 8.0) to reduce the urea concentration to less than 1M. Protein digestion was carried out by adding Trypsin (Pierce) to the tubes in the ratio of 1:30 followed by incubation overnight at 37°C. After digestion, the contents of the tube were dried in a vacuum concentrator, reconstituted in 0.1% FA, and desalted using C18 stage tips. Finally, the desalted peptides were stored at -80°C for further use.

Meningioma, Glioma and Medulloblastoma Tissue

Surgically resected tissue samples from meningioma, glioma and medulloblastoma patients were obtained from Tata Memorial Hospital, Mumbai. Representative MRI images for Meningioma, Glioma and Medulloblastoma can be found in **Figures 2A** and **4A**, respectively. All the relevant clinical information can be found in **Supplementary Table 1**. In brief, the tissue samples were flash frozen in liquid nitrogen and stored at -80°C. The protein extraction was done using Urea buffer (8M Urea, Tris-HCl buffer) with addition of Phosphatase inhibitor cocktail (Sigma Aldrich®) (21). Furthermore, the reduced and alkylated proteins were digested by addition of Trypsin (Pierce, Thermofisher Scientific) followed by overnight incubation at 37°C. The digested peptides were then vacuum dried and desalted using *in-house* C18 stage tips. The tissue sample set had Meningioma (n=6), Glioma (n=6) and Medulloblastoma (n=6) in addition to 3 control tissues for each tumor, arachnoid tissue (n=3) for Meningioma, peritumoral tissue (n=3) for Glioma and cerebellum tissue (n=3) for Medulloblastoma.

Liquid Chromatography and Triple Quadrupole Parameters

All the MRM experiments were carried out on TSQ Altis mass spectrometer (ThermoFisher Scientific, USA) coupled to a Vanquish uHPLC (ThermoFisher Scientific, USA) platform. 1µg of peptides was loaded on to a Hypersil Gold C18 column 1.9µm 100 X 2.1mm (ThermoFisher Scientific, USA) and chromatographic separation of peptides was carried out at a flow rate of 0.40ml/min for CSF samples and 0.45ml/min for tissue samples. The total time of gradient for CSF samples was 20 minutes, while that for tissue samples was 10 minutes. The buffer system was binary with Buffer A (0.1% Formic acid in water) and Buffer B (80% Acetonitrile in 0.1% Formic acid water). The gradients used have been shown in **Supplementary Table 2**. With an ESI source to the MS, the data was acquired for 20 minutes in case of CSF samples and 10 minutes in case of tissue samples. All the other MS relevant parameters (which were kept similar for both kinds of samples) are tabulated in **Supplementary Figure 1**.

BSA and MCF-7 as Quality Check Standards

To monitor the instrument behaviour and consistency in response, peptides from Bovine serum albumin (BSA) and MCF-7 cell line pellets were prepared with the urea extraction method as explained above for tissue samples. The transition lists were optimised and a final method containing 7 peptides DLGEEHFK, LVNELTEFAK, DDSPDLPK, AEFVEVTK, HLVDEPQNLIK, LGEYGFQNALIVR and QTALVELLK was prepared. Similarly, the refined list of MCF-7 had 2 proteins namely Fructose Bisphosphate Aldolase or ALDOA (GILAADESTGSIAK and ADDGRPFPQVIK) and Glyceraldehyde-3-phosphate dehydrogenase or GAPDH (LVINGNPITIFQER, GALQNIIPASTGAAK, VIPELNGK and LISWYDNEFGYSNR). To ensure the quantitative linearity of the instrument, we also injected 200ng, 400ng, 600ng, 800ng and 1µg of BSA. The results were imported in Skyline and peak areas acquired were plotted to check for linearity in the quantitative values.

Transition List Preparation and Data Analysis

Proteins and their peptides for MRM experiments were selected based on data from shotgun experiments and information available in SRM Atlas (22). The peptide sequences were imported into Skyline and transition list created with peptides having 8-20 amino acids (23). All the selected peptides were unique with 0 missed cleavage. The transition list included y-ions from "ion 3" to "last ion -3". Method files were created for the unrefined transition list for the selected proteins and initial optimization was carried out to select the best peptides and their transitions for each protein using the sample pool. After screening the transitions from the first round of experiments, a final transition file with refined transitions was prepared. Finally, the transitions of each tumour type were monitored for each sample using cycle time of 2 second and resolution of 0.7 m/z for both Q1 and Q3. The data obtained was further analysed using Skyline daily version 20.2.1.384. The fold changes were calculated using the MSstats tool in Skyline keeping a confidence interval of 95% (24). The schematic workflow of the experiment has been shown in **Figure 1A**.

RESULTS

Quality Check Using BSA and MCF-7

We performed quality checks at the following levels: 1) BSA: A single injection of 300ng of BSA was done every day when the samples were run on the instrument. The refined transition list containing 7 peptides for BSA gave uniform response. A representation of this can be seen in **Figures 1B, C**. The peak areas for BSA were found to be similar on all the five consecutive days the sample was injected. 2) The instrument response was found to be linear when increasing concentrations of BSA were injected (**Figure 1D**). 3) The CV values calculated using peak area values for 500ng of MCF-7 injected on five consecutive days were observed to be less than 15% (**Figure 1E**). These quality check steps helped us decipher the repeatability, reproducibility and efficiency of the instrument.

Monitoring of Potential Protein Markers in Meningioma and Glioma CSF

Initial optimization experiments were performed to select the best flying peptides and their transitions using pooled samples. Three peptides of 5 Proteins which includes APOA1, APOE, PTGDS, VTN and C3 were monitored for both Meningioma and Glioma samples. For the meningioma CSF samples, three peptides of APOE (SELEEQLTPVAEETR, LGPLVEQGR and AATVGSLAGQPLQER) gave a cumulative fold change of 2.21 whereas peptides of PTGDS (WFSAGLASNSSWLR, TMLLQPAGSLGSYSYR and AQGFTEDTIVFLPQTDK)

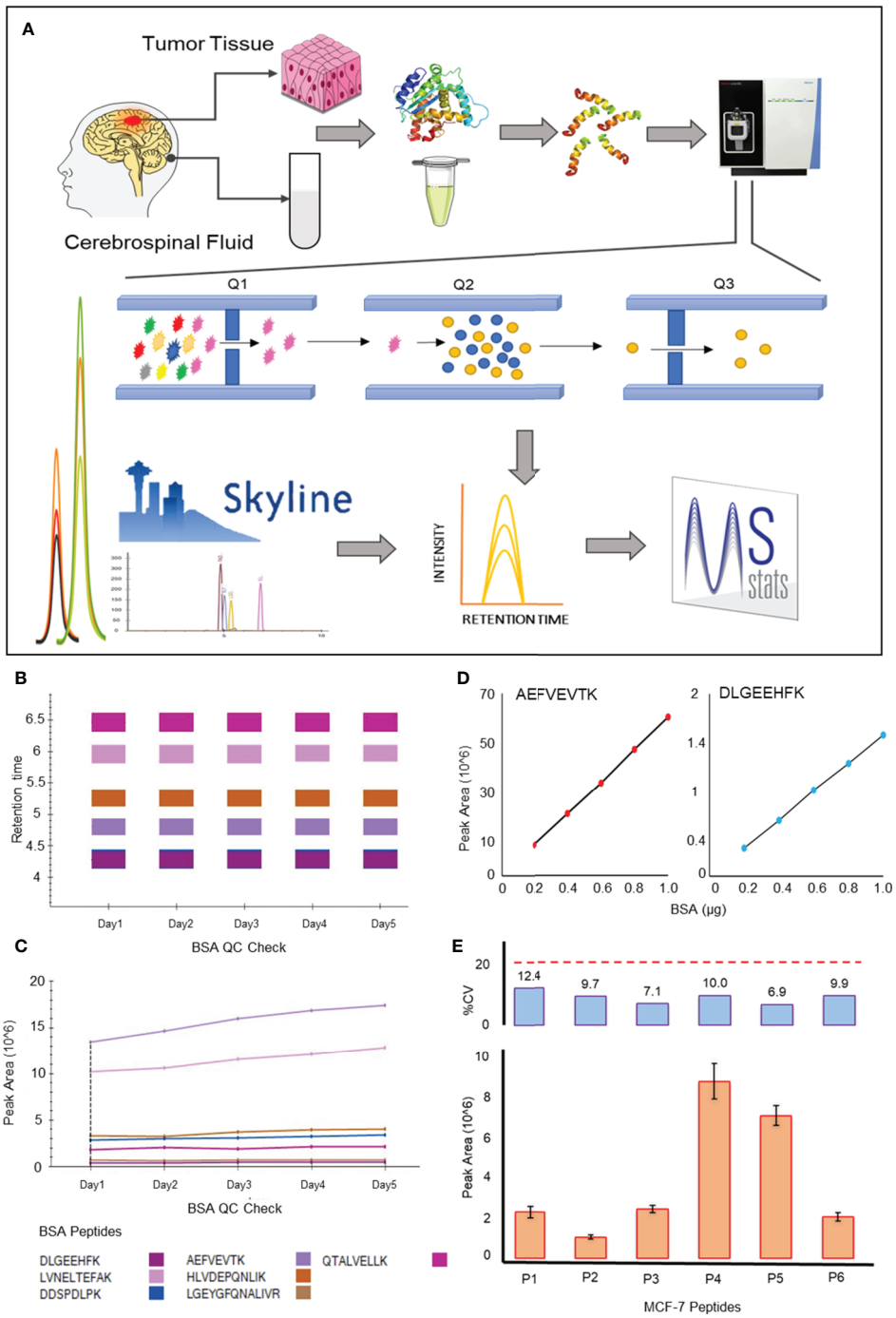


FIGURE 1 | Schematic of the MRM workflow and QC. **(A)** Schematic outline of the MRM based experiments from biological specimens; **(B, C)** Response of seven peptides of BSA (DLGEEHFK, LVNELTEFAK, DDSPDLPK, AEFVEVTK, HLVDEPQNLIK, LGEYGFQNALIVR, QTALVELLK) monitored during the experiments in terms of Retention time and Peak area respectively; **(D)**. Quantification sensitivity of the instrument using peak area against concentration (in µg) curve of two peptides of BSA injected in the increasing concentration; **(E)** shows the repeatability and variation in the response of six peptides belongs to two proteins of MCF-7 digested peptide used as a QC for the experiments (P1: GILAADESTGSIAK and P2: ADDGRPFPQVIK of ALDOA whereas P3: LVINGNPITIFQER, P4: GALQNIIPASTGAAK, P5: VIPELNGK and P6: LISWYDNEFGYSNR belongs to GAPDH).

showed a fold change of 1.52 (**Figures 2B, C**). However, the cumulative fold change of APOA1 (DYVSQFEGSALGK, LLDNWDSVTSTFSK and ATEHLSTLSEK), VTN (DVWGIE GPIDAAFTR, FEDGVLDPDYPR and SIAQYWLGCPAPGHL) and C3 (LVAYYTLIGASGQR, TGLQEVEVK and SGSDEV QVGQQR) showed upregulation in Grade I in comparison to

Grade II (**Supplementary Figure 2**). For Glioma, 3 peptides each of the above mentioned 5 proteins were chosen to look for differential expression between the grades. The same three peptides of APOE and APOA1 gave a cumulative fold change of 1.59 and 2.69 respectively in GBM when compared to low grade glioma (**Figures 3B, C**). While cumulative fold change of three peptides of VTN, PTGDS and C3 in GBM were found to be upregulated in GBM as compared to low grade glioma (**Supplementary Figure 2**). A final list of these proteins and peptides has been provided in **Supplementary Table 3**.

Monitoring of Potential Biomarkers From Tumor Tissue Specimens

Differential Protein expression for a few candidate biomarkers was observed for by comparing their expression in brain tumor tissues to that in their respective control tissues. For the meningioma tissue samples, we observed overexpression of VIM, CST3 and CLU. The three selected peptides of CLU (EILSVDCSTNNPSQAK, ELDESLQVAER and LFDSDPITVTVPVEVSR) gave a cumulative fold change of 1.41 (**Figure 2D**). While the three peptides of CST3, namely LVGGPMDASVEEEGVR, QIVAGVNYFLDVELGR and TQPNLDNCPFHDQPHLK gave a cumulative fold change of 1.53. Four peptides of VIM (SLYASSPGGVYATR, ILLAELEQLK, LGDLYEEEMR and FADLSEAANR) showed a cumulative fold change of 2.25 as compared to the arachnoid controls with a confidence interval of 95% (**Supplementary Figure 3**).

For Glioma tissue samples we screened several known biomarker proteins including ANXA1, SOD2 and VIM. Three peptides of SOD2 (GDVTAQIALQPALK, GELLEAIK and AIWNVINWENVTER) showed a cumulative positive fold change of 2.48 in Glioma samples as compared to the peritumoral controls (**Figure 3D**). ANXA1 was found to be upregulated in Glioma tissues (peptides GLGTDEDTLIEILASR, GVDEATIIDILTK and GTDVNV FNTILTTR). Apart from the above-mentioned peptides of VIM, we observed QVQSLTCEVDALK and ETNLDSLPLVDTHSK of VIM also gave good response for glioma samples. These 6 peptides showed overexpression in tumor tissue *vs* peritumoral tissue (**Supplementary Figure 3**). It is to be noted that one of the controls in this samples set, did not give good response and hence was excluded from the fold change calculations. The acquired data for this outlier has been provided in our data. Peritumoral tissues are rare to come by hence the analysis included only two controls.

For the Medulloblastoma tissue sample set, the data clearly shows overexpression of MAP2 in tumor tissue with a fold change of 1.34. Peptides TPGTPGTPSYPR, VGSLDNAHHVPGGGNVK and VDHGAEIITQSPGR were monitored for MAP2 in the individual samples (**Figure 4B**). Three peptides of SF3B2 (VGEPVALSEEER, KPGDLSDELR and YGPPPSYPNLK) showed a cumulative fold change of 2.89 whereas the four peptides of VIM (SLYASSPGGVYATR, ILLAELEQLK, LGDLYEEEMR and FADLSEAANR) gave cumulative fold change of 2.39 (**Figures 4C, D**). The final list of all the selected proteins and their peptides for Glioma, Meningioma and Medulloblastoma has been provided in **Supplementary Table 4**.

DISCUSSION

Mass Spectrometry based targeted proteomics approaches like MRM require considerable optimization and investment of time. These methods are increasingly replacing the conventional molecular biology methods owing to their superior accuracy and reproducibility over techniques like western blotting and immunohistochemistry which rely heavily on the use of antibodies. In the current study we have reported the use of MRM to accurately detect and quantify proteins and their peptides from biospecimens CSF and tumor tissue in three brain malignancies. These proteins have been reported to play important roles in development of these brain tumors and accurate detection and quantification of such proteins can greatly advance our understanding of tumor pathobiology.

The expression levels of proteins such as VTN, APOA1 and PTGDS were found to be highly up regulated in GBM as compared to low grade gliomas. Extracellular matrix (ECM) remodelling is one of mechanisms involved in tumor maturation and migration. The ECM of adult brain is characterized by absence of most of the adhesive glycoproteins that aid in cell attachment and invasion (25). However, to promote cell attachment and migration, malignant astrocytomas remodel the ECM through synthesis of VTN (26). Our MRM data from CSF of gliomas also suggests a significant up-regulation of VTN in GBMs as compared to LGGs, one of the interactors of Integrins. In CSF of Meningiomas, the levels of VTN were found to be higher in Grade I as compared to Grade II. Integrins are the cell-surface heterodimeric receptors that integrate ECM with intracellular cytoskeleton to mediate cell adhesion, migration and invasion. Malignant astrocytomas are known to express integrins $\alpha v\beta 3$ and $\alpha v\beta 5$, which bind to VTN and RGD domain of osteopontin, thereby promoting integrin-mediated cell attachment and migration (26, 27).

The protein PTGDS has been reported as a potential biomarker of meningioma in multiple studies (28, 29). Kim et.al., reported that expression of PTGDS using Western blot was found to be decreased in Meningioma CSF when compared to non-tumor controls (29). The same protein has also been reported in glioblastoma using deep learning (30). Our study validated the higher expression of PTGDS in Grade I meningioma as compared to Grade II, however the expression of this protein was found to be upregulated in GBM in comparison to LGG. Another protein, Complement C3 (C3) has been reported to be key protein in tumorigenesis of Meningiomas (31). In our study, the protein was found to be down regulated in Meningioma Grade II when compared to Grade I. The levels of this protein in gliomas were observed to be opposite to that seen in Meningiomas with the high grade GBMs showing an up-regulation over the low grade gliomas.

Our discovery dataset and literature also pointed at up-regulation of several apolipoproteins in gliomas and meningiomas (12, 31, 32). APOE plays a vital role in redistribution of intracellular lipid and tissue reconstruction in CNS through a receptor dependent pathway. Astrocytes are one of the main sites of APOE synthesis (33). Nicoll et al., showed APOE immunoreactivity in the tumor cells, macrophages and

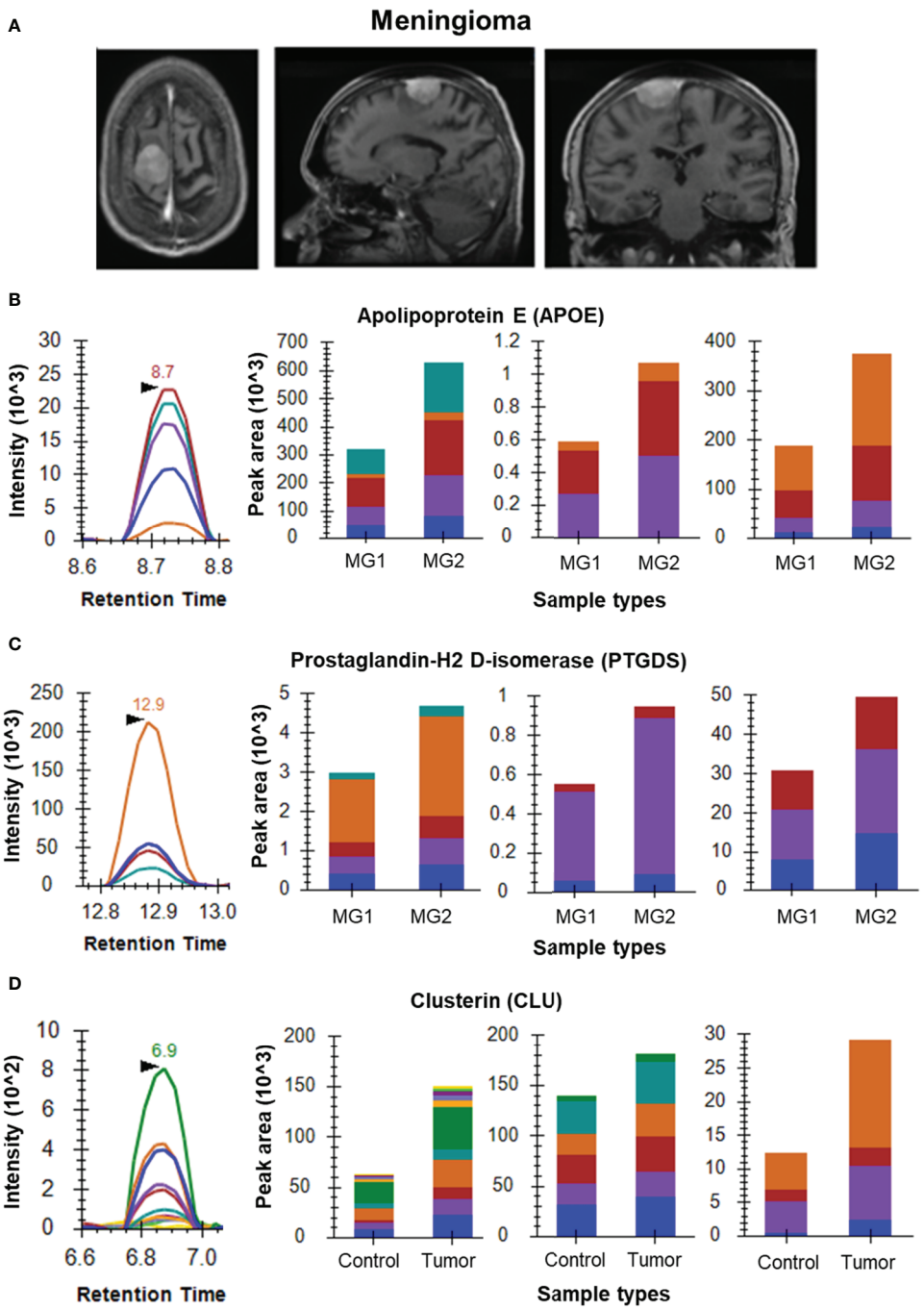


FIGURE 2 | MRM analysis of Meningioma CSF and tissue samples. **(A)** Radiological images of Meningioma – contrast MRI images (axial, sagittal and coronal views, respectively); **(B)** Representative peak shape for AATVGSLAGQPLQER and bar plots of AATVGSLAGQPLQER, LGPLVEQGR and SELEEQLTPVAEETR, respectively showing overexpression of APOE in Grade II meningioma (n=4) as compared to Grade I meningioma (n=3) in CSF samples; **(C)** Representative peak shape for AQGFTEDTIVFLPQTDK and bar plots of AQGFTEDTIVFLPQTDK, TMLLQPAGSLGSYSYR and WFSAGLASNSSWLR, respectively showing overexpression of PTGDS in Grade II meningioma as compared to Grade I meningioma in CSF samples; **(D)** Representative MRM peak for one peptide of CLU and bar plots depicting the increased expression of peptides EILSVDCSTNNPSQAK, ELDESLQVAER and LFDSDPITVTVPVEVSR of CLU in Meningioma tumor tissue samples (n=6) as compared to arachnoid controls (n=3).

nearby astrocytes, supporting the role of APOE in delivery of lipids to tumor cells and its recycling by macrophages in necrotic areas. Increased levels of lipids in serum of GBM patients and over expression of APOE in meningioma CSF sample in comparison to non-tumor CSF have also been reported (34) and (29). In the current study, we have observed an upregulation

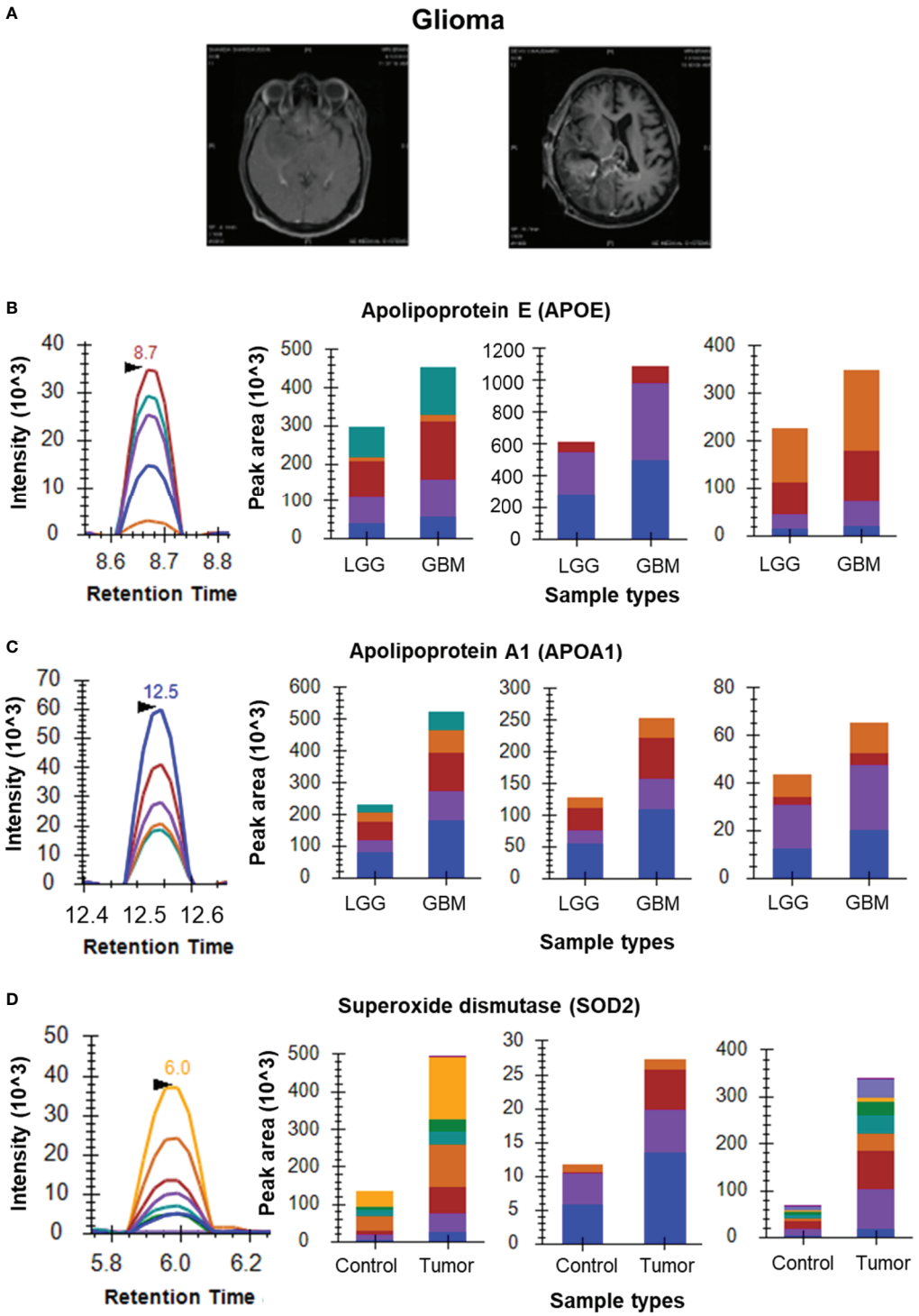


FIGURE 3 | MRM analysis of Glioma CSF and tissue samples. **(A)** T1 contrast images showing Low Grade Glioma and High Grade Glioma, respectively; **(B)** Representative peak shape for AATVGSLAGQPLQER and bar plots of AATVGSLAGQPLQER, LGPLVEQGR and SELEEQLTPVAEETR, respectively showing overexpression of APOE in CSF samples of GBM (n=5) as compared to LGG (n=5); **(C)** Representative peak shape for LLDNWDSVTSTFSK and bar plots of LLDNWDSVTSTFSK, DYVSQFEGSALGK and ATEHLSTLSEK, respectively showing overexpression of APOA1 in GBM CSF samples as compared to LGG CSF samples; **(D)** Representative MRM peak for one peptide of SOD2 and bar plots depicting the increased expression of peptides GDVTAQIALQPALK, LLDNWDSVTSTFSK and ATEHLSTLSEK of SOD2 in Glioma tumor tissue samples (n=6) as compared to peritumoral controls (n=3).

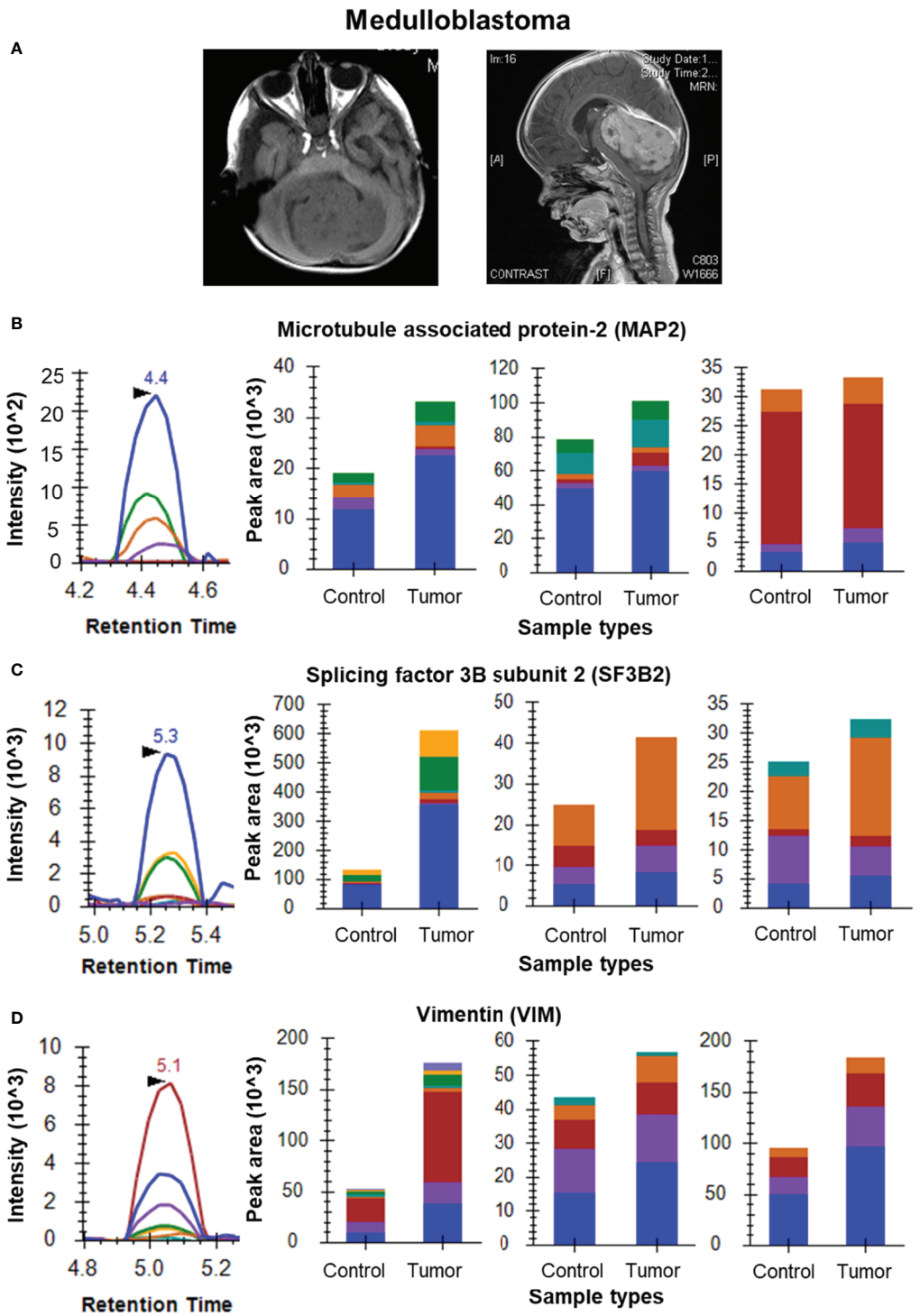


FIGURE 4 | MRM analysis of Medulloblastoma tissue samples. **(A)** MRI images of a medulloblastoma showing a mass on T1 weighted image and the sagittal contrast image showing the extent of the tumor, respectively; **(B)** MRM peak shape for TPGTPGTPSYPR of MAP2 and bar plots of TPGTPGTPSYPR, VGSLDNAHHVPGGGNVK and VDHGAEIITQSPGR, respectively showing overexpression of MAP2 in tumor tissue (n=6) as compared to cerebellar controls (n=3); **(C)** MRM peak shape for VGEPVALSEEER of SF3B2 and bar plots of VGEPVALSEEER, KPGDLSDELR and YGPPPSYPNLK, respectively showing overexpression of SF3B2 in tumor tissue as compared to cerebellar controls; **(D)** MRM peak shape for SLYASSPGGVYATR of VIM and bar plots of SLYASSPGGVYATR, ILLAELEQLK and FADLSEAANR, respectively showing overexpression of VIM in tumor tissue as compared to cerebellar controls.

of APOE and APOA1 in GBMs when compared to LGGs and upregulation of APOE and downregulation of APOA1 in Meningioma Grade II tissue compared to Meningioma Grade I.

For tissue samples, the MRM experiments were performed on proteins with roles in tumor pathology curated from available literature. The Human Protein Atlas reports a higher expression of VIM in Gliomas (35, 36). Mukherjee et al. validated the expression of VIM in Meningiomas using MRM (37). In the current study, we have reported four peptides for VIM, not reported elsewhere further strengthening the claim for its use as a biomarker for meningiomas as reported by Mukherjee et al. In Meningiomas, the protein CLU was found to be upregulated when compared to controls. It has been reported as a meningioma associated marker in the literature and known to inhibit apoptosis (31, 38). We have also observed an up regulation of CST3 in meningioma tumour samples when compared to control tissues. CST3 is an inhibitor of cysteine proteases and has been reported to have a positive alteration in high-grade meningiomas (39).

Many studies have reported the role of ANXA1 and SOD2 in gliomas, by virtue of their increased expression levels in the tumor tissues (40). ANXA1 is known to alter regular cell proliferation, differentiation and works as a substrate of EGFR. We observed that the levels of peptides for ANXA1 and SOD2 were upregulated in glioma tissues in comparison to peritumoral control tissues.

MAP2 is a well-known neural marker for medulloblastoma as reported in a few studies (41–43). It is a frequently considered marker during IHC of MB samples. It has been observed in MB tumors irrespective of the age of the patient (44). Our data shows the overexpression of this protein in MB tumors as compared to controls. Cancerous mutations in splicing factor SF3B2 have been reported to affect the ubiquitinylation pathways and hence associated with cancer. SF3B2 has been reported as a potential gene marker in many diseases including medulloblastoma in DisGeNET and has appeared as one of the significantly dysregulated proteins in our discovery dataset (45). One of our recent studies on medulloblastoma has also highlighted the significance of splicing events in medulloblastoma disease biology (10). In our data, VIM was observed to be overexpressed in medulloblastoma as compared to controls in accordance with literature (12, 46, 47).

In conclusion, our study highlights the importance of targeted proteomics in detection and validation of proteins with roles in pathobiology of brain tumors. From our MRM experiments using CSF we report that APOE could be a potential tumor progression marker in Meningioma and Glioma. The trends for APOA1 were found to be opposite in Gliomas and Meningiomas, with higher expression in GBMs and Grade I meningiomas as compared to the LGGs and Grade II meningiomas, respectively. We also observed upregulation of VTN, PTGDS and APOA1 in CSF of GBM patients in comparison to CSF of LGG patients. The protein Vimentin was observed to be overexpressed in all the three brain tumor tissue samples. We have also validated tumor markers such as CLU and CST3 for Meningioma, ANXA1 and SOD2 in Glioma and MAP2 for Medulloblastoma. With the advancements in gene sequencing techniques, routine diagnosis for complex cancers has become easier, faster and efficient. However, there remains a greater reliance on the age-old molecular biology methods of Immunohistochemistry (IHC) and Fluorescence in-situ hybridization (FISH). In addition to the non-specific binding towards proteins, these antibody-based methods also have notoriously high chances of inter-observer variability leading to differences in grading the tumors (48). The MRM technique owing to its quantitative accuracy and sensitivity, can offer a suitable alternative to the more labour-intensive molecular biology techniques currently used. The development of an analytical method and assay based on MRM involves a multitude of aspects such as generating a calibration curve, determination of analytical specificity (selectivity or interference), analytical sensitivity, carryover, precision, recovery of assay, matrix effect, recovery of immunoprecipitation, dilution integrity, stability, reproducibility, and quality control (QC) of samples (49). Successful validation of proteins on a large cohort with easily obtainable biospecimens from patients can pave way to designing panels of protein markers with ability to distinguish between the grades of these tumors, thus providing a faster and more accurate alternative to the existing modalities.

DATA AVAILABILITY STATEMENT

The targeted proteomic datasets are deposited in the Panorama Public and can be accessed through this link: https://panoramaweb.org/Brain_Tumor_marker_MRM.url

ETHICS STATEMENT

The studies involving Glioma patients was approved by the Institutional Review Board (IRB) at Tata Memorial Hospital (TMC-ACTREC Project No.15). The study on meningioma patients was approved as part of an institutional review board at TMH (TMC-ACTREC project no.149). For use of samples from patients with Medulloblastomas the approval was obtained from TMH IRB (TMC-ACTREC project no.197). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SS, SG, MP, and DB planned the experiment. AM, and PS acquired clinical samples for the study. SE performed the grading and histopathological studies for the acquired clinical samples. DB, MP, SM, SG, NG, and MK carried out experiments. DB, MP, and SG compiled the data and wrote the manuscript. MN was involved in

instrument setup and data acquisition. DB, MP, and AS performed statistical analysis. All authors contribted to the article and approved the submitted version.

FUNDING

The study was funded through MHRD-UAY Project UCHHATAR AVISHKAR YOJANA; UAY-(MHRD), project #34_IITB (2016) to SS, MASSFIIT (Mass Spectrometry Facility, IIT Bombay; BT/PR13114/INF/22/206/2015) for MS-based proteomics experiments.

ACKNOWLEDGMENTS

We would like to thank MHRD for funding the project and MASSFIITB (Mass Spectrometry Facility, IIT Bombay) for the technical support. We would also like to thank Dr. Kishore Gollapalli and Ms. Susmita Ghosh for their help in this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.548243/full#supplementary-material

REFERENCES

1. Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK, et al. Connecting Genomic Alterations to Cancer Biology With Proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discovery* (2013) 3:1108–12. doi: 10.1158/2159-8290.CD-13-0219
2. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an Immeasurable Source of Knowledge. *Wspolczesna Onkol* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136
3. Omenn GS. Advances of the HUPO Human Proteome Project With Broad Applications for Life Sciences Research. *Expert Rev Proteomics* (2017) 14:109–11. doi: 10.1080/14789450.2017.1270763
4. Iwamoto N, Shimada T. Recent Advances in Mass Spectrometry-Based Approaches for Proteomics and Biologics: Great Contribution for Developing Therapeutic Antibodies. *Pharmacol Ther* (2018) 185:147–54. doi: 10.1016/j.pharmthera.2017.12.007
5. Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, et al. Glioma. *Nat Rev Dis Prim* (2015) 1:15017. doi: 10.1038/nrdp.2015.17
6. Ostrom QT, Gittleman H, Liao P, Rouse C, Chen Y, Dowling J, et al. CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007–2011. *Neuro Oncol* (2014) 16 Suppl 4:iv1–63. doi: 10.1093/neuonc/nou223
7. Ostrom QT, Gittleman H, Liao P, Vecchione-Koval T, Wolinsky Y, Kruchko C, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2010-2014. *Neuro Oncol* (2017) 19:v1– v88. doi: 10.1093/neuonc/nox158
8. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, et al. Molecular Subgroups of Medulloblastoma: the Current Consensus. *Acta Neuropathol* (2012) 123:465–72. doi: 10.1007/s00401-011-0922-z
9. Boulos PT, Dumont AS, Mandell JW, Jane JA. Meningiomas of the Orbit: Contemporary Considerations. *Neurosurg Focus* (2001) 10:1–10. doi: 10.3171/foc.2001.10.5.6
10. Kp M, Kumar A, Biswas D, Moiyadi A, Shetty P , Gupta T, et al. The Proteomic Analysis Shows Enrichment of RNA Surveillance Pathways in Adult SHH and Extensive Metabolic Reprogramming in Group 3 Medulloblastomas. *Brain Tumor Pathol* (2021) 1:3. doi: 10.1007/s10014-020-00391-x
11. Gahoi N, Malhotra D, Moiyadi A, Varma SG, Gandhi MN, Srivastava S. Multi-Pronged Proteomic Analysis to Study the Glioma Pathobiology Using Cerebrospinal Fluid Samples. *Proteomics - Clin Appl* (2018) 12:e1700056. doi: 10.1002/prca.201700056
12. Sharma S, Ray S, Moiyadi A, Sridhar E, Srivastava S. Quantitative Proteomic Analysis of Meningiomas for the Identification of Surrogate Protein Markers. *Sci Rep* (2014) 4:7140. doi: 10.1038/srep07140
13. Method of the Year 2012. *Nat Methods* (2013) 10:1. doi: 10.1038/nmeth.2329
14. Shi T, Song E, Nie S, Rodland KD, Liu T, Qian WJ, et al. Advances in Targeted Proteomics and Applications to Biomedical Research. *Proteomics* (2016) 16:2160–82. doi: 10.1002/pmic.201500449
15. Faria SS, Morris CFM, Silva AR, Fonseca MP, Forget P, Castro MS, et al. A Timely Shift From Shotgun to Targeted Proteomics and How it can be Groundbreaking for Cancer Research. *Front Oncol* (2017) 7:13. doi: 10.3389/fonc.2017.00013
16. Kumar V, Ray S, Aggarwal S, Biswas D, Jadhav M, Yadav R, et al. Multiplexed Quantitative Proteomics Provides Mechanistic Cues for Malaria Severity and Complexity. *Commun Biol* (2020) 3:683. doi: 10.1038/s42003-020-01384-4
17. Ishizaki J, Takemori A, Suemori K, Matsumoto T, Akita Y, Sada K, et al. Targeted Proteomics Reveals Promising Biomarkers of Disease Activity and Organ Involvement in Antineutrophil Cytoplasmic Antibody-Associated Vasculitis. *Arthritis Res Ther* (2017) 19:218. doi: 10.1186/s13075-017-1429-3
18. Cardozo KHM, Lebkuchen A, Okai GG, Schuch RA, Viana LG, Olive AN, et al. Establishing a Mass Spectrometry-Based System for Rapid Detection of SARS-Cov-2 in Large Clinical Sample Cohorts. *Nat Commun* (2020) 11:6201. doi: 10.1038/s41467-020-19925-0
19. Song E, Gao Y, Wu C, Shi T, Nie S, Fillmore TL, et al. Targeted Proteomic Assays for Quantitation of Proteins Identified by Proteogenomic Analysis of Ovarian Cancer. *Sci Data* (2017) 4:170091. doi: 10.1038/sdata.2017.91
20. Addona T, Abbatiello S, Carr SA. From Skepticism to Embrace: the Role of Targeted Mass Spectrometry in Validating Proteomics. *Clin Chem* (2020) 66:973–4. doi: 10.1093/clinchem/hvaa111
21. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell* (2016) 166:755–65. doi: 10.1016/j.cell.2016.05.069
22. Kusebauch U, Campbell DS, Deutsch EW, Chu CS, Spicer DA, Brusniak MY, et al. Human Srmatlas: a Resource of Targeted Assays to Quantify the Complete Human Proteome. *Cell* (2016) 166:766–78. doi: 10.1016/j.cell.2016.06.041
23. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* (2010) 26:966–8. doi: 10.1093/bioinformatics/btq054
24. Choi M, Chang CY, Clough T, Broudy D, Killeen T, MacLean B, et al. Msstats: an R Package for Statistical Analysis of Quantitative Mass Spectrometry-Based Proteomic Experiments. *Bioinformatics* (2014) 30:2524–6. doi: 10.1093/bioinformatics/btu305
25. Gladson CL, Cheresh DA. Glioblastoma Expression of Vitronectin and the αvβ3 Integrin: Adhesion Mechanism for Transformed Glial Cells. *J Clin Invest* (1991) 88:1924–32. doi: 10.1172/JCI115516
26. Gladson CL, Wilcox JN, Sanders L, Gillespie GY, Cheresh DA. Cerebral Microenvironment Influences Expression of the Vitronectin Gene in Astrocytic Tumors. *J Cell Sci* (1995) 108: (Pt 3):947–56.
27. Denhardt DT, Noda M. Osteopontin Expression and Function: Role in Bone Remodeling. *J Cell Biochem* (1998) 30–31:92–102. doi: 10.1002/(sici)1097-4644(1998)72:30/31+<92::aid-jcb13>3.3.co;2-1
28. Aarhus M, Lund-Johansen M, Knappskog PM. Gene Expression Profiling of Meningiomas: Current Status After a Decade of Microarray-Based Transcriptomic Studies. *Acta Neurochir (Wien)* (2011) 153:447–56. doi: 10.1007/s00701-010-0906-0
29. Kim JH, Lee SK, Yoo YC, Park NH, Park DB, Yoo JS, et al. Proteome Analysis of Human Cerebrospinal Fluid as a Diagnostic Biomarker in Patients With Meningioma. *Med Sci Monit* (2012) 18:BR450–60. doi: 10.12659/MSM.883538
30. Wong KK, Rostomily R, Wong STC. Prognostic Gene Discovery in Glioblastoma Patients Using Deep Learning. *Cancers (Basel)* (2019) 11. doi: 10.3390/cancers11010053

31. Abbritti RV, Polito F, Cucinotta M, Giudice C, Caffo M, Tomasello C, et al. Meningiomas and Proteomics: Focus on New Potential Biomarkers and Molecular Pathways. *Cancer Genomics Proteomics* (2016) 13:369–79.
32. Hashemi M, Pooladi M, Abad SKR. The Investigation of Changes in Proteins Expression (Apolipoprotein A1 and Albumin) in Malignant Astrocytoma Brain Tumor. *J Cancer Res Ther* (2014) 10:107–11. doi: 10.4103/0973-1482.131413
33. McCarron MO, Nicoll JAR, Stewart J, Ironside JW, Mann DMA, Love S, et al. The Apolipoprotein E E2 Allele and the Pathological Features in Cerebral Amyloid Angiopathy-Related Hemorrhage. *J Neuropathol Exp Neurol* (1999) 58:711–8. doi: 10.1097/00005072-199907000-00005
34. Gollapalli K, Ray S, Srivastava R, Renu D, Singh P, Dhali S, et al. Investigation of Serum Proteome Alterations in Human Glioblastoma Multiforme. *Proteomics* (2012) 12:2378–90. doi: 10.1002/pmic.201200002
35. Schiffer D, Giordana MT, Mauro A, Migheli A, Germano I, Giaccone G. Immunohistochemical Demonstration of Vimentin in Human Cerebral Tumors. *Acta Neuropathol* (1986) 70:209–19. doi: 10.1007/BF00686074
36. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas - a Tool for Pathology. *J Pathol* (2008) 216:387–93. doi: 10.1002/path.2440
37. Mukherjee S, Biswas D, Gadre R, Jain P, Syed N, Stylianou J, et al. Comprehending Meningioma Signaling Cascades Using Multipronged Proteomics Approaches & Targeted Validation of Potential Markers. *Front Oncol* (2020) 10:1600. doi: 10.3389/fonc.2020.01600
38. Zhang H, Kim JK, Edwards CA, Xu Z, Taichman R, Wang CY. Clusterin Inhibits Apoptosis by Interacting With Activated Bax. *Nat Cell Biol* (2005) 7:909–15. doi: 10.1038/ncb1291
39. Dunn J, Lenis VP, Hilton DA, Warta R, Herold-Mende C, Hanemann CO, et al. Integration and Comparison of Transcriptomic and Proteomic Data for Meningioma. *Cancers (Basel)* (2020) 12. doi: 10.3390/cancers12113270
40. Gollapalli K, Ghantasala S, Atak A, Rapole S, Moiyadi A, Epari S, et al. Tissue Proteome Analysis of Different Grades of Human Gliomas Provides Major Cues for Glioma Pathogenesis. *Omi A J Integr Biol* (2017) 21:275–84. doi: 10.1089/omi.2017.0028
41. Pietsch T, Haberler C. Update on the Integrated Histopathological and Genetic Classification of Medulloblastoma - a Practical Diagnostic Guideline. *Clin Neuropathol* (2016) 35:344–52. doi: 10.5414/NP300999
42. Gupta T, Sarkar C, Rajshekhar V, Chatterjee S, Shirsat N, Muzumdar D, et al. Indian Society of Neuro-Oncology Consensus Guidelines for the Contemporary Management of Medulloblastoma. *Neurol India* (2017) 65:315–32. doi: 10.4103/0028-3886.201841
43. Joshi P, Katsushima K, Zhou R, Meoded A, Stapleton S, Jallo G, et al. The Therapeutic and Diagnostic Potential of Regulatory Noncoding Rnas in Medulloblastoma. *Neuro-Oncol Adv* (2019) 1:vdz023. doi: 10.1093/noajnl/vdz023
44. Giordana MT, Cavalla P, Dutto A, Borsotti L, Chiò A, Schiffer D. Is Medulloblastoma the Same Tumor in Children and Adults? *J Neurooncol* (1997) 35:169–76. doi: 10.1023/A:1005762727257
45. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The Disgenet Knowledge Platform for Disease Genomics: 2019 Update. *Nucleic Acids Res* (2020) 48:D845–55. doi: 10.1093/nar/gkz1021
46. Seok JY, Kim SH, Lee YH, Kwon J, Kim TS. Vimentin and Survivin Expression Rates as Prognostic Factors in Medulloblastoma. *KOREAN J Pathol* (2007) 41:87–94.
47. Cruz-Sanchez FF, Rossi ML, Hughes JT, Esiri MM, Coakham HB. Medulloblastoma - an Immunohistological Study of 50 Cases. *Acta Neuropathol* (1989) 79:205–10. doi: 10.1007/BF00294380
48. Anagnostou VK, Welsh AW, Giltnane JM, Siddiqui S, Lioeaga C, Gustavson M, et al. Analytic Variability in Immunohistochemistry Biomarker Studies. *Cancer Epidemiol Biomarkers Prev* (2010) 19:982–91. doi: 10.1158/1055-9965.EPI-10-0097
49. Kim H, Sohn A, Yeo I, Yu SJ, Yoon JH, Kim Y. Clinical Assay for AFP-L3 by Using Multiple Reaction Monitoring–Mass Spectrometry for Diagnosing Hepatocellular Carcinoma. *Clin Chem* (2018) 64:1230–8. doi: 10.1373/clinchem.2018.289702

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 Ghantasala, Pai, Biswas, Gahoi, Mukherjee, KP, Nissa, Srivastava, Epari, Shetty, Moiyadi and Srivastava. 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 Oncology

ORIGINAL RESEARCH
published: 01 June 2021
doi: 10.3389/fonc.2021.576326

Proteomic Characterization of Cytoplasmic Lipid Droplets in Human Metastatic Breast Cancer Cells

Alyssa S. Zembroski[†], *Chaylen Andolino*[†], *Kimberly K. Buhman* and *Dorothy Teegarden*[*]

Department of Nutrition Science, Purdue University, West Lafayette, IN, United States

OPEN ACCESS

Edited by:
Bhaswati Chatterjee, National Institute of Pharmaceutical Education and Research, India

Reviewed by:
Fuquan Yang, Institute of Biophysics (CAS), China
Daniele Vergara, University of Salento, Italy

***Correspondence:**
Dorothy Teegarden
dteegard@purdue.edu

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *25 June 2020*
Accepted: *10 May 2021*
Published: *01 June 2021*

Citation:
Zembroski AS, Andolino C, Buhman KK and Teegarden D (2021) Proteomic Characterization of Cytoplasmic Lipid Droplets in Human Metastatic Breast Cancer Cells. Front. Oncol. 11:576326. doi: 10.3389/fonc.2021.576326

One of the characteristic features of metastatic breast cancer is increased cellular storage of neutral lipid in cytoplasmic lipid droplets (CLDs). CLD accumulation is associated with increased cancer aggressiveness, suggesting CLDs contribute to metastasis. However, how CLDs contribute to metastasis is not clear. CLDs are composed of a neutral lipid core, a phospholipid monolayer, and associated proteins. Proteins that associate with CLDs regulate both cellular and CLD metabolism; however, the proteome of CLDs in metastatic breast cancer and how these proteins may contribute to breast cancer progression is unknown. Therefore, the purpose of this study was to identify the proteome and assess the characteristics of CLDs in the MCF10CA1a human metastatic breast cancer cell line. Utilizing shotgun proteomics, we identified over 1500 proteins involved in a variety of cellular processes in the isolated CLD fraction. Interestingly, unlike other cell lines such as adipocytes or enterocytes, the most enriched protein categories were involved in cellular processes outside of lipid metabolism. For example, cell-cell adhesion was the most enriched category of proteins identified, and many of these proteins have been implicated in breast cancer metastasis. In addition, we characterized CLD size and area in MCF10CA1a cells using transmission electron microscopy. Our results provide a hypothesis-generating list of potential players in breast cancer progression and offers a new perspective on the role of CLDs in cancer.

Keywords: breast cancer, metastasis, cytoplasmic lipid droplets, proteomics, triacylglycerol

INTRODUCTION

Breast cancer is the leading type of cancer among women in the United States and is predicted to account for 30% of new cancer cases in 2021 (1). Although breast cancer survival rates are relatively high if the cancer remains localized, life expectancy dramatically decreases once the cancer metastasizes to distant organs such as bone and lung (2). Therefore, understanding the characteristic features of metastatic breast cancer cells is critical in order to develop strategies to prevent the progression of breast cancer.

Metastatic breast cancer cells often exhibit altered lipid metabolism, which is an adaptation that allows cells to survive in nutrient-depleted conditions (3). One of these alterations includes the accumulation of neutral lipid in cytoplasmic lipid droplets (CLDs). The degree of CLD accumulation associates with breast cancer aggressiveness (4–6); however, the mechanisms behind this relationship are incompletely understood. CLDs are composed of a neutral lipid core of triacylglycerol (TAG) and/or cholesteryl esters surrounded by a phospholipid monolayer and associated proteins (7). The role of CLDs differs depending on cell type, for example serving as the body's TAG storage pool in adipocytes (8), acting as a local energy source for skeletal and cardiomyocytes (9, 10), and mediating the process of dietary fat absorption in enterocytes (11). Although multiple hypotheses exist for how CLDs contribute to cancer progression, including protection from cellular stress or serving as a storage pool for fatty acids that can be used for cellular energy, biosynthetic processes, or signaling (6, 12, 13), the exact role of CLDs in metastatic breast cancer cells has not been determined.

Proteins that associate with CLDs serve a variety of functions, but their role in metastasis in unknown. A common function of CLD proteins is to mediate TAG synthesis and lipolysis, reflecting the main purpose of CLDs in storing neutral lipid and maintaining cellular lipid homeostasis (14). However, recent functional studies of CLD proteins demonstrate novel cellular roles for CLDs by regulating cellular protein location, degradation, and functional activity. For example, histone proteins and transcription factors sequester at the CLD as a mechanism to regulate gene expression (15–17). In addition, some CLD proteins are destined for degradation (18) such as apolipoprotein B-100, a component of very-low-density lipoproteins, which translocates from the endoplasmic reticulum (ER) to the CLD for degradation in hepatocytes (19, 20). Finally, CLD proteins may also have specific functions on the CLD, for example mediating inflammatory signaling pathways (21–23). Despite the identification of proteins involved in a variety of roles in CLD proteomic studies, the functional significance of the majority of CLD proteins has yet to be uncovered. Further, the functional significance of CLD proteins in metastatic breast cancer cells and whether they reflect unique roles for CLDs in cancer is unknown.

The purpose of this study was to identify the proteome of CLDs in metastatic breast cancer cells to generate hypotheses about how CLDs promote breast cancer progression and contribute to altered lipid metabolism and/or other cell functions. To do this, we performed untargeted shotgun proteomic analysis and utilized transmission electron microscopy (TEM) to identify the proteome and characteristics of CLDs from the human metastatic breast cancer cell line, MCF10CA1a.

Abbreviations: ABHD5, 1-acylglycerol-3-phosphate O-acyltransferase ABHD5; ATGL, adipose triglyceride lipase; CANX, calnexin; CHO, Chinese Hamster Ovary; CLD, cytoplasmic lipid droplet; EMT, epithelial-mesenchymal transition; ER, endoplasmic reticulum; FF, floating fraction; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GO_BP, Gene Ontology Biological Process; GPAT4, glycerol-3-phosphate acyltransferase 4; HMGCS1, hydroxymethylglutaryl-CoA synthase; IQGAP1, IQ motif containing GTPase activating protein 1; LC-MS/MS, liquid chromatography-tandem mass spectrometry; LFQ, label-free quantification; MS, mass spectrometry; NSDHL, sterol-4-alpha-carboxylate 3-dehydrogenase; PBS, phosphate buffered saline; PCYT1A, choline-phosphate cytidylyltransferase A; PLIN, perilipin; PNPLA2, patatin-like phospholipase domain-containing protein 2; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SQLE, squalene epoxidase; STK24, serine/threonine kinase 24; TAG, triacylglycerol; TEM, transmission electron microscopy; WCL, whole cell lysate.

MATERIALS AND METHODS

Cell Culture

The MCF10CA1a human metastatic mammary cell line was utilized for these studies. Cells were cultured in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12, 1:1), supplemented with 5% horse serum, 100 units/mL penicillin, and 100 µg/mL streptomycin in a humidified environment at 37°C with 5% CO_2. Cells were harvested at 70-80% confluence for each experiment.

CLD Isolation

Cells from eight 150 mm dishes were pooled and considered one sample. Four samples were prepared for CLD isolation as follows. Cells were rinsed with phosphate buffered saline (PBS, pH 7.4, 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4) scraped and pelleted by centrifugation. CLDs were isolated from pelleted cells using a previously established sucrose gradient ultracentrifugation protocol (24, 25). Briefly, cells were lysed in ice cold sucrose lysis buffer (175 mM sucrose, 10 mM HEPES and 1 mM EDTA pH 7.4) and disrupted by passing through a 23 gauge, 1 inch needle. An aliquot was taken representing the whole cell lysate (WCL) to be used for later applications. The remaining lysate was transferred into a 13.2 mL Open-Top Thinwall UltraClear tube (Beckman Coulter, #344059) and ice-cold lysis buffer was layered on top of the lysate. Samples were centrifuged at 100,000 x g at 4°C for one hour. After centrifugation, the white floating fraction (FF) from each sample was aspirated using a pipette. The remaining soluble and pellet fractions were removed in 1 mL increments. Samples were stored at -80°C until analysis.

Triacylglycerol and Protein Concentration

TAG concentrations of each fraction were measured using the Wako L-Type Triglyceride M kit (FUJIFILM Wako Diagnostics U.S.A.).

Protein concentrations of each fraction were measured using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific).

Validation of CLD Isolation by Western Blotting

An aliquot of each isolated fraction (CLD fraction through pellet fraction) and the WCL was delipidated with 2:1 chloroform: methanol, then proteins were precipitated with ice-cold acetone. The precipitated proteins were pelleted by centrifugation, then dried and resuspended in Laemmli loading buffer. Samples were subjected to SDS-PAGE using a 12% Tris-glycine gel (Bio-Rad #4561046). Samples were loaded into the gel by volume: 10 µL each of the FF through fraction 10, 5 µL of the pellet fraction and 5 µL of the WCL. See **Supplementary Figure 1** for representative Ponceau stain demonstrating differences in protein levels between fractions. The membrane was probed with one of the following primary antibodies at a 1:1,000 concentration (PLIN3, Sigma-Aldrich HPA006427; GAPDH, Cell Signaling Technologies #14C10; CANX, Santa Cruz Biotechnology SC-11397). After washing, a fluorescent secondary antibody was added at a concentration of 1:10,000 (LI-COR IRDye donkey anti-rabbit 680RD, 926-68073). Membranes were imaged using the LI-COR Odyssey CLx Imaging System (LI-COR Biosciences).

Transmission Electron Microscopy

One 60 mm dish of MCF10CA1a cells and one 60 mm dish of MCF10A-*ras* cells were prepared for TEM. Cells were fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer, rinsed, and embedded in agarose. Small pieces of cell pellet were post-fixed in 1% osmium tetroxide containing 0.8% potassium ferricyanide and stained in 1% uranyl acetate. They were then dehydrated with a graded series of ethanol, transferred into acetonitrile, and embedded in EMbed-812 resin. Thin sections were cut on a Reichert-Jung Ultracut E ultramicrotome and post-stained with 4% uranyl acetate and lead citrate. Images were acquired on a FEI Tecnai T12 electron microscope equipped with a tungsten source and operating at 80kV.

CLD Size Analysis

Acquired TEM images were analyzed using ImageJ (26). 50 cells were counted and used for CLD analysis. CLD diameter was assessed using ImageJ.

Immunocytochemistry

MCF10CA1a cells were cultured in a #1.5H-N high performance glass bottom 12 well plate (Cellvis) and processed for immunofluorescence microscopy. The cells were fixed in 4% paraformaldehyde, permeabilized with 0.1% Triton X-100, and blocked with BlockAid (Invitrogen, B10710). Cells were probed with antibodies for PLIN3, SQLE, and NSDHL (Sigma, HPA006427; SantaCruz Biotechnologies, sc-271651; Atlas Antibodies, HPA000571, respectively). Proteins were detected using secondary AlexaFluor antibodies (Life Technologies, A-21070 and A-21052), and cells were counterstained for neutral lipids using 1 µg/mL 4,4-difluoro-1,3,5,7,8-pentamethyl-4-bora-3a,4a-diaza-s-indacene (BODIPY 493/503; Life Technologies, Grand Island, NY, United States), and for nuclei using 300 nM 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI; Invitrogen, D1306). Samples were imaged using a Nikon A1R-MP inverted confocal microscope (Nikon Instruments Inc., Melville, NY, United States). Images were acquired using the Plan Apo λ 100x Oil objective, 76.63 µm pinhole size, and DAPI, FITC, and Cy5 lasers. All image processing was conducted using Nikon NIS-Elements AR acquisition and analysis software. A Landweber 2D deconvolution algorithm was implemented, with point scan confocal modality, clear noise, and 12, 12, 12 iterations.

CLD Protein Isolation and In-Solution Digestion

An aliquot of each CLD fraction containing 50 µg protein was prepared for proteomic analysis. The CLD fractions were delipidated and precipitated as above. The dried protein pellets were reduced and solubilized using 10 mM dithiothreitol/8 M urea, then alkylated using iodoethanol. Samples were dried using a vacuum centrifuge. Proteins were digested with 4 µg Trypsin/Lys-C Mix, Mass Spec Grade (Promega) per sample using a barocycler at 50°C, 20 kpsi, 60 cycles (Barocycler NEP2320, Pressure Biosciences, INC). Peptides were cleaned with MacroSpin C18 spin columns (The Nest Group, Inc) and dried using a vacuum centrifuge. Dried peptides were resuspended in 3% acetonitrile/0.1% formic acid in preparation for mass spectrometry.

Liquid Chromatography/Tandem Mass Spectrometry (LC-MS/MS)

Samples were analyzed by reverse-phase LC-ESI-MS/MS system using the Dionex UltiMate 3000 RSLC nano System coupled to the Orbitrap Fusion Lumos Mass Spectrometer (Thermo Fisher Scientific). Peptides were loaded onto a trap column (300 mm ID ´ 5 mm) packed with 5 mm 100 Å PepMap C18 medium, then separated on a reverse phase column (50-cm long × 75 µm ID) packed with 2 µm 100 Å PepMap C18 silica (Thermo Fisher Scientific). The column temperature was maintained at 50°C. All MS measurements were performed in positive ion mode using a 120 minute LC gradient and standard data-dependent mode. MS data were acquired with a Top20 data-dependent MS/MS scan method.

LC-MS/MS Data Analysis

LC-MS/MS data were analyzed using MaxQuant software version 1.6.3.4 (27–29). Data was searched against the UniProtKB *Homo sapiens* reference proteome (www.uniprot.org). Trypsin/P and Lys-C were selected with a maximum of 2 missed cleavages. Oxidation of methionine was set as a variable modification, iodoethanol set as a fixed modification. First search peptide mass tolerance was set to 20 ppm, main search peptide mass tolerance was set to 10 ppm. False discovery rate was set to 1%. Match between runs was selected and Label-free quantification (LFQ) was used.

Proteomic Data Analysis

Reverse identifications and contaminants were removed from the dataset. LFQ values were subjected to Log2 transformation. A protein was considered identified if it was present in at least three out of four samples. Uniprot accession numbers in the Majority

Protein IDs column were used to categorize proteins into Gene Ontology Biological Process (GO_BP) terms using The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (30, 31). Functional relationships between proteins were visualized using STRING version 11 (32).

RESULTS

Characterization of CLDs in MCF10CA1a Cells

To characterize CLDs in the metastatic breast cancer MCF10CA1a cell line, we visualized cells by TEM. A representative MCF10CA1a cell containing CLDs is shown in **Figure 1A**. CLDs present within the cell are highlighted (**Figure 1B**). To determine the distribution of CLDs across cells, we assessed the number and diameter of CLDs per cell (**Table 1** and **Figure 2**). Ninety percent of cells counted contained CLDs, and the number of CLDs per cell ranged from 0-41. CLD diameter also varied across cells. CLD diameter ranged from 0.17-1.38 μm (**Figure 2**), with an average CLD diameter of 0.58 μm. As expected, only 10% of non-metastatic MCF10A-*ras* cells of the same cell series analyzed contained CLDs (data not shown). A representative MCF10A-*ras* cell without CLDs is shown in **Supplementary Figure 2**. Due to the absence of CLDs in most MCF10A-*ras* cells, we were unable to isolate CLDs from MCF10A-ras cells and therefore only assessed the proteome of CLDs from MCF10CA1a cells.

TABLE 1 | Number and size of CLDs within MCF10CA1a cells. 50 cells were counted and used for the analysis.

% of cells containing CLDs	# CLDs per cell	Average # CLDs per cell	CLD diameter range (μm)	Average CLD diameter (μm)
90	0-41	12	0.17-1.38	0.58

*CLD diameter was measured using ImageJ. Distribution of CLD diameters is shown in **Figure 2**.*

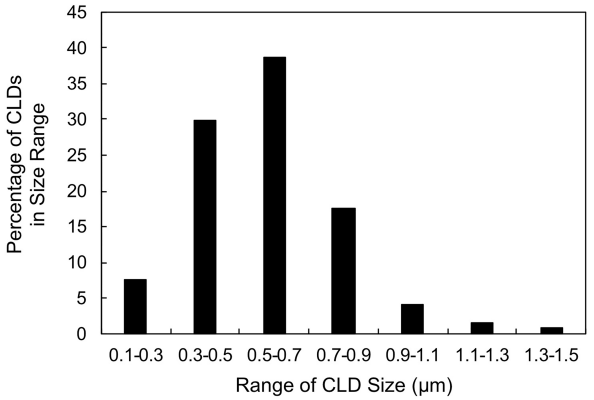


FIGURE 2 | Cytoplasmic lipid droplet (CLD) size distribution. Percentage of CLDs analyzed in **Table 1** within the indicated size range. 50 cells were counted and used for the analysis. CLD diameter was measured using ImageJ.

CLD Isolation From MCF10CA1a Cells

To confirm successful isolation of CLDs from MCF10CA1a cells, we determined the TAG to protein ratio of each isolated fraction after sucrose density gradient ultracentrifugation (**Figure 3A**). A high TAG to protein ratio in the floating fraction (FF) indicates the presence of CLDs. In addition, we determined the purity of our isolation based on the presence of specific cell component markers in each isolated fraction (**Figure 3B**). Perilipin (PLIN) 3, a bona-fide CLD-associated protein and marker of CLDs (33), is present in the FF. PLIN3 resides in the cytosol but associates with CLDs when CLDs are present (34), which is consistent with its identification in the soluble fractions. The localization of PLIN3

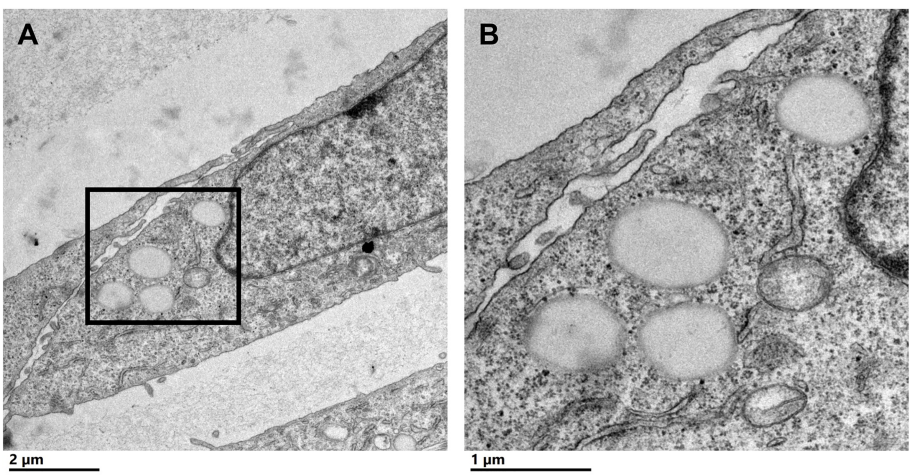


FIGURE 1 | Cytoplasmic lipid droplets (CLDs) are present in MCF10CA1a cells. **(A)** Representative transmission electron microscopy (TEM) image of a MCF10CA1a cell containing CLDs (boxed region), scale bar 2 μm. **(B)** Magnified image of the CLDs present in **(A)** scale bar 1 μm.

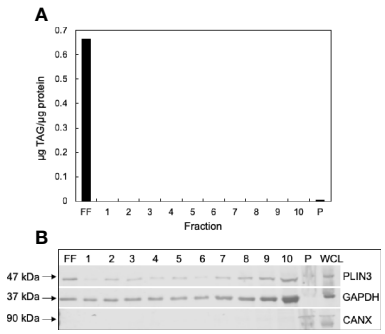


FIGURE 3 | Validation of cytoplasmic lipid droplet (CLD) isolation. **(A)** Triacylglycerol (TAG) to protein ratio of each isolated fraction. CLDs were isolated from MCF10CA1a cells using sucrose density gradient ultracentrifugation. Fractions were removed sequentially from the top of the gradient to the bottom. Floating fraction (FF): isolated CLDs, 1-10: soluble fractions, P: pellet. **(B)** Western blot of isolated fractions and whole cell lysate (WCL). Fractions were loaded by volume: 10 μL FF-10, 5 μL P and WCL. Membrane was probed for markers of CLDs (PLIN3), cytosol (GAPDH), and endoplasmic reticulum (CANX). Approximate molecular weight markers for each protein are listed. See **Supplementary Figure 1** for a representative Ponceau stain reflecting the relative levels of protein in each fraction.

to CLDs was confirmed by immunocytochemistry (**Figure 4**). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a cytosolic marker, is present in the FF and soluble fractions but absent in the pellet fraction (**Figure 3B**). GAPDH is identified in CLD proteomic studies of certain cell types (35, 36), and the identification of GAPDH in the FF suggests GAPDH is a CLD-associated protein in MCF10CA1a cells. Calnexin (CANX), a marker of ER, is present in only the pellet fraction (**Figure 3B**), as expected based on published CLD isolation protocols (24). Isolated fractions were loaded by volume and therefore contain different amounts of protein; see representative Ponceau stain in **Supplementary Figure 1** for the relative amount of protein in each fraction.

Proteomic Characterization of CLDs in MCF10CA1a Cells

To determine the proteome of CLDs in MCF10CA1a cells, we performed untargeted shotgun proteomic analysis of the isolated CLD fraction using LC-MS/MS. We identified 1534 proteins (**Supplementary Table 1**) that are involved in a wide array of cellular functions (**Figure 5A**). Many of the proteins identified have functions in DNA and RNA metabolic processes (19%) and protein metabolism (18%). To determine whether a specific category of proteins was overrepresented in our dataset, we sorted proteins by Gene Ontology Biological Process (GO_BP) enrichment (**Figure 5B**). Cell-cell adhesion was the most enriched category of proteins identified, followed by translational initiation, and cotranslational protein targeting to membrane (**Figure 5B** and **Supplementary Table 2**). Surprisingly, proteins involved in lipid metabolism comprise only 3% of the proteins identified (**Figure 5A**), and lipid metabolic terms are not represented within the top 50 most enriched GO_BP categories (**Supplementary Table 2**). Low abundance of lipid metabolism proteins is in contrast to other CLD proteomic studies, where they are frequently enriched (14). We analyzed the 41 proteins identified as associated with lipid metabolism (**Figure 6**). Most of these proteins are involved in cholesterol synthesis, including hydroxymethylglutaryl-CoA synthase, cytoplasmic (HMGCS1), squalene monooxygenase (SQLE), and sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating (NSDHL). The localization of SQLE and NSDHL to CLDs was confirmed by immunocytochemistry (**Figure 7**). Both SQLE and NSDHL are shown to concentrate around CLDs. Other identified proteins have basic roles in CLD metabolism, including lipolysis [patatin-like phospholipase domain-containing protein 2/adipose triglyceride lipase (PNPLA2/ATGL), 1-acylglycerol-3-phosphate O-acyltransferase ABHD5 (ABHD5)], phospholipid synthesis [choline-phosphate cytidylyltransferase A (PCYT1A)], TAG synthesis [glycerol-3-phosphate acyltransferase 4 (GPAT4)], and the PLINs (PLIN3 and PLIN4).

Proteins Involved in Cell-Cell Adhesion Are Implicated in Breast Cancer Progression

We further analyzed the proteins belonging to the cell-cell adhesion category, as this was the most enriched GO_BP term of proteins identified (**Figure 5B**). To determine how CLDs and their proteins contribute to breast cancer metastasis, we chose proteins in the cell-cell adhesion category that also had GO_BP terms in cell migration and signaling. Proteins with these criteria are listed in **Table 2**. Many of these proteins have been shown to

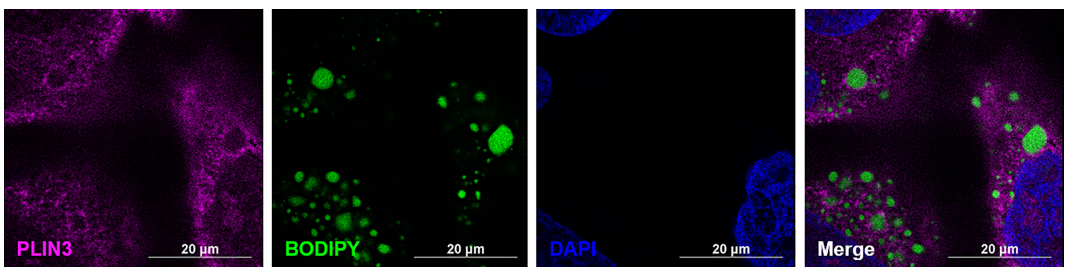


FIGURE 4 | PLIN3 surrounds cytoplasmic lipid droplets (CLDs) in MCF10CA1a cells. Representative immunofluorescence images of MCF10CA1a cells. Cells were stained with Alexa Fluor 633 to visualize PLIN3, BODIPY to visualize CLDs, and DAPI to visualize nuclei. Signals from all three channels were merged for the final image.

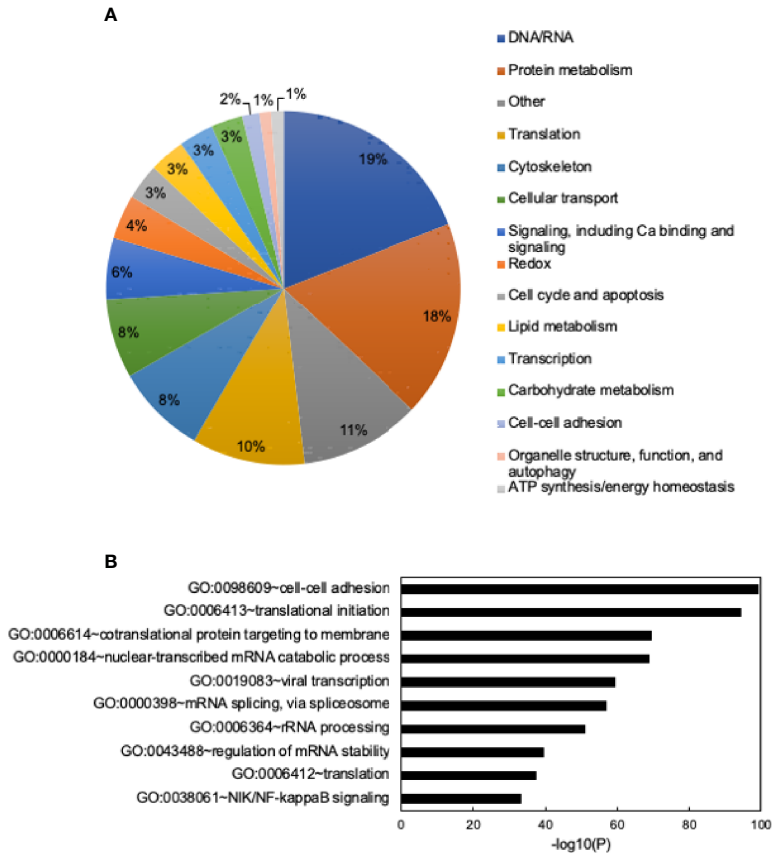


FIGURE 5 | General functions of identified proteins and Gene Ontology (GO) term enrichment. **(A)** Identified proteins grouped into general categories. Data shown as a percent of total proteins identified. Categories with the highest to lowest percent of proteins listed from top to bottom and are read clockwise around the pie chart. **(B)** Chart of the top 10 most enriched Gene Ontology Biological Process (GO_BP) terms. Most to least enriched term listed from top to bottom. Data shown as -log10 (p-value). Enrichment scores/p-values calculated in DAVID. See **Supplementary Table 2** for full list of enriched GO terms and enrichment scores.

promote breast cancer progression, and the references for each are included in **Table 2**.

DISCUSSION

To determine mechanisms by which CLDs contribute to breast cancer metastasis, we examined the characteristics and proteome of CLDs in the human metastatic breast cancer cell line, MCF10CA1a, using TEM and LC-MS/MS. We found that the majority of MCF10CA1a cells analyzed contain multiple CLDs that associate with a variety of proteins. To our knowledge, this is the first report of the proteome of CLDs in metastatic breast cancer cells. We identified 1534 proteins in the isolated CLD fraction representing a wide array of cellular functions. Many of the proteins identified are implicated in breast cancer metastasis. Our results provide a hypothesis-generating list of potential players contributing to cancer progression and provide a new perspective on the role of CLDs in metastatic breast cancer.

Our results are consistent with previous work demonstrating that neutral lipid accumulation in breast cancer cells correlates with cancer aggressiveness (4, 5, 78–80). MCF10CA1a cells are the most metastatic in the MCF10A series of breast cancer progression (81) and contain twelve times more TAG than the non-metastatic MCF10A-*ras* cell line from which they were derived (82). Consistently, while most MCF10CA1a cells analyzed contained at least one CLD (**Table 1**), almost no CLDs were present in non-metastatic MCF10A-*ras* cells (**Supplementary Figure 2**). The underlying mechanism driving increased CLDs in metastatic MCF10CA1a cells and not in MCF10A-*ras* cells is not clear, however, several factors may contribute. For example, metastatic breast cancer cells may have an increased ability, compared to non-metastatic cells, to take up or synthesize FA and cholesterol which are used as substrates for TAG and cholesteryl ester synthesis and subsequently stored in CLDs (6). Overall, these results support our use of the MCF10CA1a cell line as a model of mammary metastasis to investigate the CLD proteome.

CLD size is often used to estimate the amount of cellular neutral lipid storage and the metabolic state of the cell. For example, cells that store large amounts of TAG, such as adipocytes (83) and enterocytes (84), have large CLDs (ranging up to 100 μm), whereas

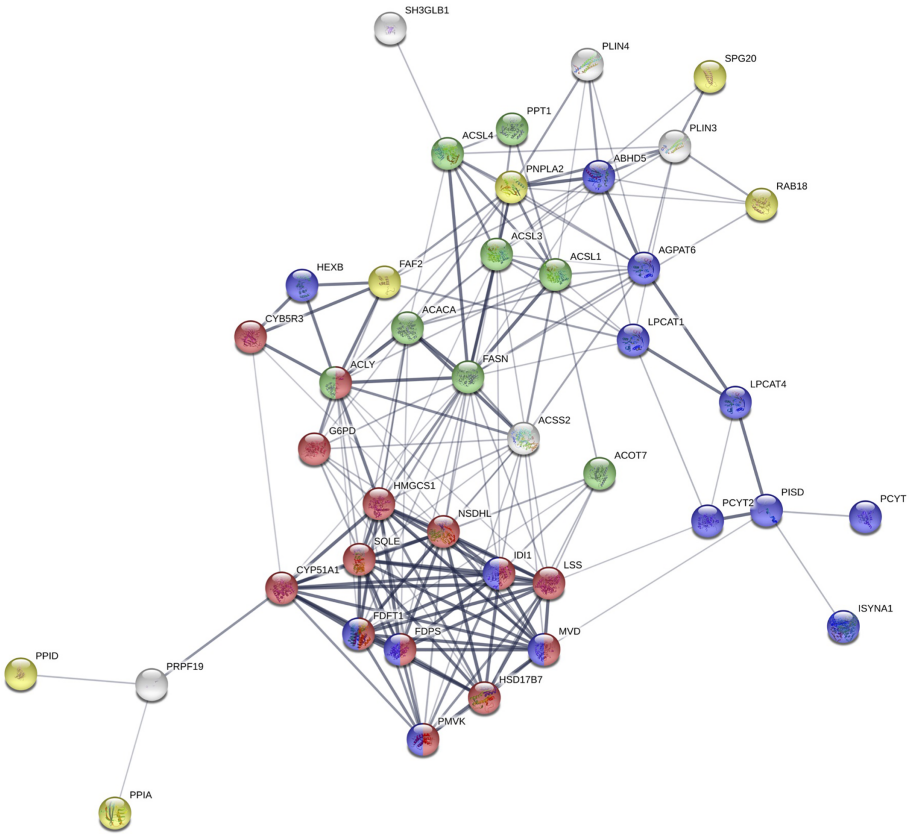


FIGURE 6 | STRING analysis of identified proteins involved in lipid metabolism. Proteins with known functions in lipid metabolism and those associated with lipid-related Gene Ontology Biological Process (GO_BP) terms. Red: cholesterol biosynthetic process; green: fatty-acyl-CoA metabolic process; purple: phospholipid metabolic process; yellow: lipid droplet organization.

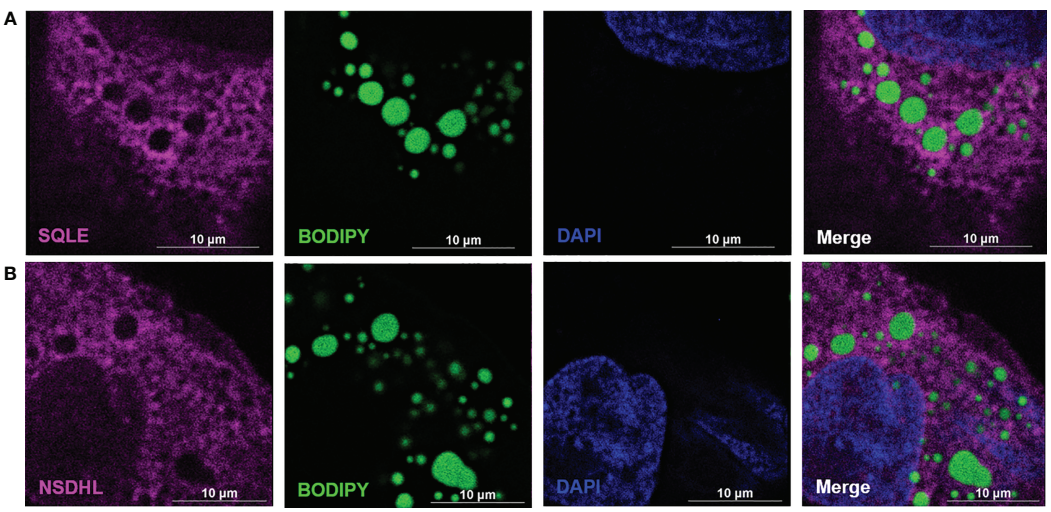


FIGURE 7 | SQLE and NSDHL localize to cytoplasmic lipid droplets (CLDs) in MCF10CA1a cells. Representative immunofluorescence images of MCF10CA1a cells. Cells were stained with Alexa Fluor 633 to visualize SQLE **(A)** and NSDHL **(B)**, BODIPY to visualize CLDs, and DAPI to visualize nuclei. Signals from all three channels were merged for the final image in **(A, B)**.

TABLE 2 | Proteins in cell-cell adhesion are associated with breast cancer metastasis.

Protein name	Gene	General function	Proposed role in breast cancer metastasis	References
IQ motif containing GTPase activating protein 1	IQGAP1	Scaffold protein; signaling and cytoskeleton dynamics	Promotes cell proliferation, migration, tumor growth	(37–41)
Serine/threonine kinase 24	STK24	MAPK signaling	Promotes cell proliferation, tumor growth	(42)
S100 calcium binding protein P	S100P	Calcium signaling	Promotes cell proliferation, migration, motility	(43–45)
Fascin actin-bundling protein 1	FSCN1	Actin-binding protein; cell adhesion, motility, migration	Promotes metastasis through NFκB and STAT3 signaling	(46–51)
GIPC PDZ domain containing family member 1	GIPC1	Scaffold protein; signaling	Involved in cell cycle, apoptosis, motility	(52, 53)
Profilin 1	PFN1	Actin-binding protein; cytoskeletal dynamics	Suppresses cell migration and cell cycle	(54–57)
Tumor-associated calcium signal transducer 2	TACSTD2	Calcium signaling	Promotes cell growth, migration, proliferation through AKT signaling	(58, 59)
Syndecan binding protein	SDCBP	Adaptor protein; signaling and cytoskeletal dynamics	Promotes cell proliferation, growth, motility, cell cycle	(60–63)
RAB1A, member RAS oncogene family	RAB1A	Vesicle trafficking from ER to Golgi	Involved in cell proliferation, migration, EMT; involved in mTORC1 signaling	(64, 65)
STE20 like kinase	SLK	Apoptosis, cytoskeletal dynamics	Promotes cell migration	(66)
Coronin 1B	CORO1B	Actin-binding protein; cell motility	Involved in cell cycle progression	(67)
Heat shock protein family A (Hsp70) member 5	HSPA5	Protein folding	Promotes cell motility and proliferation	(68–70)
Microtubule associated protein RP/EB family member 1	MAPRE1	Microtubule dynamics	Promotes cell proliferation and tumor growth	(71)
Radixin	RDX	Binds actin	Involved in cell motility; interacts with ERBB2 receptors	(72, 73)
Signal transducer and activator of transcription 1	STAT1	Transcription factor; responds to cytokines and growth factors	Either promotes or inhibits tumor growth	(74–77)

List of proteins within the Gene Ontology Biological Process (GO_BP) category "cell-cell adhesion" that also have GO_BP terms of signaling and/or cell migration.

other cell types tend to have smaller CLDs. Consistent with the size of CLDs in cell types that do not store large amounts of TAG, including skeletal myocytes (85), hepatocytes (86), and Chinese Hamster Ovary (CHO) cells (87), the diameter of CLDs in the MCF10CA1a cells averaged 0.58 μm (**Table 1**). Further, the distribution of CLDs of various sizes in MCF10CA1a cells (**Figure 2**) may reflect different pools of CLDs that have potentially distinct functions (35, 88). For example, specific pools of CLDs in brown adipose tissue are differentially involved in fatty acid oxidation or TAG synthesis (89). It is possible that unique pools of CLDs with different functions may exist in MCF10CA1a cells; however, this requires further investigation.

The proteome of CLDs identified in MCF10CA1a cells has similarities and differences compared to that of other cell types. Many of the proteins identified are consistent with the general categories of proteins commonly found on CLDs. These include proteins involved in lipid and CLD metabolism, translation, protein folding and degradation, cytoskeleton, and histones (14). Several of the proteins identified involved in lipid metabolism have been validated as CLD-associated proteins and also have functional roles at the CLD surface, including PLIN3 in CLD maintenance (90), GPAT4 (91) and PCYT1A (92, 93) in CLD expansion and size, ATGL in CLD lipolysis (94), and NSDHL in cholesterol synthesis (95, 96). The identification of lipid metabolism proteins on CLDs in MCF10CA1a cells suggests CLDs across cell types may share similar lipid metabolic machinery and core CLD proteins.

A key difference between the proteome of CLDs in MCF10CA1a cells and that of other cell types is the representation of proteins in the commonly identified categories. For example, lipid metabolism was not a highly enriched protein category in MCF10CA1a cells as it is in other cell types (14). Further, many of the proteins we identified in the lipid metabolism category are involved in cholesterol metabolism, suggesting CLDs in MCF10CA1a cells may store cholesterol (97, 98). Consistently, cholesteryl ester accumulation and altered cholesterol metabolism is a common feature of cancer (99, 100). We found that two enzymes involved in cholesterol synthesis, NSDHL and SQLE, concentrate in areas around CLDs in MCF10CA1a cells (**Figure 7**).

The identification of NSDHL with CLDs in breast cancer cells is consistent with previous observations of its functional association with CLDs and role in metastasis. NSDHL modifies lanosterol before its synthesis into cholesterol (101), and has been shown to localize to CLDs upon oleate loading in CHO cells (95) and in COS-7 cells (96). In fact, oleate loading and CLD formation in CHO cells decreased the synthesis of C-27 sterols, which includes cholesterol, and increased the synthesis of precursor sterols, including lanosterol (95). These results suggest the localization of NSDHL to CLDs may be a mechanism to regulate cholesterol synthesis. NSDHL has also been shown to promote breast cancer progression. NSDHL is present at higher protein levels in metastatic compared to non-metastatic breast cancer cell lines (102), and knockdown of NSDHL in metastatic BT-20 and MDA-MB-231 cells reduced cell viability, colony formation, and cell migration (102). However, whether these effects are due to lack of NSDHL itself or lack of cholesterol synthesis due to NSDHL inhibition is unclear. Thus, the localization of NSDHL to CLDs in MCF10CA1a cells shown in this study suggests that it may promote breast cancer progression by regulating cholesterol synthesis. Future studies are required to determine the role of NSDHL on CLDs in MCF10CA1a cells and its contribution to metastasis.

The identification of SQLE with CLDs in breast cancer cells is also consistent with previous observations of its functional

association with CLDs and role in metastasis. SQLE catalyzes the epoxidation of squalene and is considered the second rate-limiting step in cholesterol synthesis (103). SQLE localizes to CLDs in yeast cells (104) and has been shown to regulate CLD dynamics. For example, inhibition of SQLE results in CLD clustering and squalene accumulation in yeast (105), and CLD accumulation in MCF7 breast cancer cells (106). SQLE may regulate CLD dynamics by interacting with microprotein CASIMO1 (106). CASIMO1 in MCF7 cells was shown to regulate the expression of SQLE as well as CLD formation. How CASIMO1 and/or SQLE influences CLDs is not clear; however, it may involve changes in the cytoskeleton. Interestingly, SQLE has been identified as an oncogene in breast cancer cells (107), suggesting it plays a role in breast cancer metabolism. Consistently, inhibiting SQLE in MCF7 breast cancer cells reduces cell proliferation and ERK phosphorylation/activation (106), which is a key factor involved in initiating cell proliferation and migration in cancer cells (108). ERK phosphorylation and activation has previously been shown to be regulated by SQLE in other cell types including hepatocellular carcinoma cells (109) and lung squamous cell carcinoma cells (110). In fact, SQLE-mediated cholesterol synthesis preserves breast cancer stem cell stemness through PI3K/AKT signaling, another proliferative survival pathway, upon stabilization of SQLE mRNA by long non-coding RNA 030 and poly(rC) binding protein 2 (111). Therefore, the metabolites produced by the action of SQLE may activate cell signaling pathways necessary for cancer cell proliferation. Overall, these results suggest that the localization of enzymes involved in cholesterol synthesis to CLDs in MCF10CA1a cells may be a metabolic adaptation by cancer cells that stimulates cell proliferation. Future studies are required to determine the role of SQLE on CLDs in MCF10CA1a cells.

Instead of lipid metabolism proteins representing the majority of the CLD proteome, proteins with roles in cell-cell adhesion, translation, and mRNA metabolism were the most prevalent in the CLD fraction of MCF10CA1a cells, suggesting these proteins may have a novel functional role on CLDs in cancer. The most enriched category of proteins identified were those involved in cell-cell adhesion. This is particularly interesting, since loss of cell adhesion is a critical first step in the metastatic cascade (112). Many of the proteins identified in this category have been implicated in breast cancer metastasis (**Table 2**), suggesting CLDs may play a novel role in this process. For example, CLDs may serve as a hub for signaling pathways and cytoskeletal remodeling proteins that are needed to facilitate the epithelial-mesenchymal transition (EMT). However, CLD proteins may either play an active role at the CLD surface or may be mislocalized from their typical cell location, which could interrupt their function and contribute to metastasis. Future studies are required to determine the role of signaling and cytoskeletal proteins identified in **Table 2** on CLDs in MCF10CA1a cells.

Another category of proteins identified in the isolated CLD fraction of MCF10CA1a cells is RNA binding proteins and translational proteins. Some of these proteins are also implicated in cell motility and breast cancer metastasis (113, 114), suggesting their localization on CLDs contributes to metastatic potential. For example, downregulation of the RNA-binding protein ZBP1 in metastatic breast cancer cells increased cell migration by altering the expression of mRNAs involved in cell-cell adhesion, cytoskeleton, and cell proliferation (115). In addition, overexpression of the 60S ribosomal subunit RPL15 in circulating tumor cells isolated from patients with metastatic breast cancer increased the translation of ribosomal proteins and proteins involved in cell proliferation, and when injected into mice resulted in increased metastasis and tumor formation (116). Interestingly, RNA localizes to CLDs in human mast cells (117) and ribosomes localize to CLDs in human monocyte U937 cells and leukocytes (118). It is possible that CLDs in MCF10CA1a cells house RNA-binding and translational proteins in order to facilitate localized gene expression and protein translation to promote cell migration. This hypothesis requires testing in future experiments.

Validation of proteins identified in the CLD fraction by methods such as immunocytochemistry is needed to conclude that a protein associates with CLDs. It is possible that some proteins identified localize near, but may not directly associate with, CLDs. CLDs interact with multiple cellular organelles (119) and proteins associated with an interacting organelle may be isolated with the CLD fraction. Since we have not validated all the proteins in our analysis for cellular location *via* another mechanism, only hypotheses about their localization and function in cancer progression can be made. Despite this limitation, our analysis has generated a novel list of proteins that can be studied in more detail in future experiments.

In summary, we characterized CLDs and the CLD proteome isolated from the human metastatic breast cancer cell line, MCF10CA1a. The identification of an interesting variety of proteins in the isolated CLD fraction reflects both similarities with CLDs in other cell types, as well as differences that may support a novel role of CLDs in cancer. It is possible that proteins associated with CLDs in metastatic cancer cells may play a role in permitting the advantageous metabolic plasticity that supports cancer progression. It would be interesting to assess the similarities and differences of CLD proteomes in other metastatic breast cancer cell lines which may further our understanding of cancer progression and identify factors that can be targeted to prevent metastasis. In conclusion, this study provides a new perspective on the role of CLDs in breast cancer metastasis.

DATA AVAILABILITY STATEMENT

Raw LC-MS/MS data is available on the Mass Spectrometry Interactive Virtual Environment (MassIVE) data repository at ftp://massive.ucsd.edu/MSV000086731/.

AUTHOR CONTRIBUTIONS

KB and DT conceived and designed the study. AZ and CA performed the experiments, analyzed the data, and wrote the first

draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Purdue University Center for Cancer Research, Indiana Clinical Translational Science Institute NIH/NCRR Grant #TR000006, and the National Institutes of Health 5R01CA232589.

ACKNOWLEDGMENTS

The Purdue Proteomics Facility and the Purdue Imaging Facility in the Bindley Bioscience Center was utilized for this study. In addition, we thank Laurie Mueller and Bob Seiler of the Purdue Life Science Microscopy Facility for assistance with sample preparation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.576326/full#supplementary-material

Supplementary Figure 1 | Representative Ponceau stain for Western blots. Fractions were loaded by volume: 10 μL floating fraction (FF)-10, 5 μL pellet (P) and whole cell lysate (WCL). Membrane demonstrates the relative amount of protein per lane.

Supplementary Figure 2 | CLDs are not present in non-metastatic MCF10A-*ras* cells. Representative transmission electron microscopy image (TEM) of a MCF10A-*ras* cell, scale bar 5 μm.

Supplementary Table 1 | LC-MS/MS results of all proteins identified.

Supplementary Table 2 | Gene Ontology_Biological Process (GO_BP) enrichment. Sheet 1: GO_BP enrichment of identified proteins; Sheet 2: Individual identified proteins and their GO_BP terms.

REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654
2. Jin X, Mu P. Targeting Breast Cancer Metastasis. *Breast Cancer (Auckl)* (2015) 9(Suppl 1):23–34. doi: 10.4137/BCBCR.S25460
3. Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid Metabolic Reprogramming in Cancer Cells. *Oncogenesis* (2016) 5:e189. doi: 10.1038/oncsis.2015.49
4. Nieva C, Marro M, Santana-Codina N, Rao S, Petrov D, Sierra A. The Lipid Phenotype of Breast Cancer Cells Characterized by Raman Microspectroscopy: Towards a Stratification of Malignancy. *PloS One* (2012) 7(10):e46456. doi: 10.1371/journal.pone.0046456
5. Abramczyk H, Surmacki J, Kopec M, Olejnik AK, Lubecka-Pietruszewska K, Fabianowska-Majewska K. The Role of Lipid Droplets and Adipocytes in Cancer. Raman Imaging of Cell Cultures: MCF10A, MCF7, and MDA-MB-231 Compared to Adipocytes in Cancerous Human Breast Tissue. *Analyst* (2015) 140(7):2224–35. doi: 10.1039/C4AN01875C
6. Shyu P Jr., Wong XFA, Crasta K, Thibault G. Dropping in on Lipid Droplets: Insights Into Cellular Stress and Cancer. *Biosci Rep* (2018) 38(5):e20180764. doi: 10.1042/BSR20180764
7. Walther TC, Chung J, Farese RV Jr. Lipid Droplet Biogenesis. *Annu Rev Cell Dev Biol* (2017) 33:491–510. doi: 10.1146/annurev-cellbio-100616-060608
8. Qi Y, Sun L, Yang H. Lipid Droplet Growth and Adipocyte Development: Mechanistically Distinct Processes Connected by Phospholipids. *Biochim Biophys Acta Mol Cell Biol Lipids* (2017) 1862(10 Pt B):1273–83. doi: 10.1016/j.bbalip.2017.06.016
9. Bosma M. Lipid Droplet Dynamics in Skeletal Muscle. *Exp Cell Res* (2016) 340(2):180–6. doi: 10.1016/j.yexcr.2015.10.023
10. Goldberg IJ, Reue K, Abumrad NA, Bickel PE, Cohen S, Fisher EA, et al. Deciphering the Role of Lipid Droplets in Cardiovascular Disease: A Report From the 2017 National Heart, Lung, and Blood Institute Workshop. *Circulation* (2018) 138(3):305–15. doi: 10.1161/CIRCULATIONAHA.118.033704
11. D'Aquila T, Hung YH, Carreiro A, Buhman KK. Recent Discoveries on Absorption of Dietary Fat: Presence, Synthesis, and Metabolism of Cytoplasmic Lipid Droplets Within Enterocytes. *Biochim Biophys Acta* (2016) 1861(8 Pt A):730–47. doi: 10.1016/j.bbalip.2016.04.012
12. Cruz ALS, Barreto EA, Fazolini NPB, Viola JPB, Bozza PT. Lipid Droplets: Platforms With Multiple Functions in Cancer Hallmarks. *Cell Death Dis* (2020) 11(2):105. doi: 10.1038/s41419-020-2297-3
13. Petan T, Jarc E, Jusovic M. Lipid Droplets in Cancer: Guardians of Fat in a Stressful World. *Molecules* (2018) 23(8):1941. doi: 10.3390/molecules 23081941
14. Zhang C, Liu P. The New Face of the Lipid Droplet: Lipid Droplet Proteins. *Proteomics* (2019) 19(10):e1700223. doi: 10.1002/pmic.201700223
15. Mejhert N, Kuruvilla L, Gabriel KR, Elliott SD, Guie MA, Wang H, et al. Partitioning of MLX-Family Transcription Factors to Lipid Droplets Regulates Metabolic Gene Expression. *Mol Cell* (2020) 77(6):1251–64. e9. doi: 10.1016/j.molcel.2020.01.014
16. Johnson MR, Stephenson RA, Ghaemmaghami S, Welte MA. Developmentally Regulated H2Av Buffering Via Dynamic Sequestration to Lipid Droplets in Drosophila Embryos. *Elife* (2018) 7:e36021. doi: 10.7554/eLife.36021
17. Li Z, Thiel K, Thul PJ, Beller M, Kuhnlein RP, Welte MA. Lipid Droplets Control the Maternal Histone Supply of Drosophila Embryos. *Curr Biol* (2012) 22(22):2104–13. doi: 10.1016/j.cub.2012.09.018
18. Bersuker K, Olzmann JA. Establishing the Lipid Droplet Proteome: Mechanisms of Lipid Droplet Protein Targeting and Degradation. *Biochim Biophys Acta Mol Cell Biol Lipids* (2017) 1862(10 Pt B):1166–77. doi: 10.1016/j.bbalip.2017.06.006
19. Ohsaki Y, Cheng J, Fujita A, Tokumoto T, Fujimoto T. Cytoplasmic Lipid Droplets Are Sites of Convergence of Proteasomal and Autophagic Degradation of Apolipoprotein B. *Mol Biol Cell* (2006) 17(6):2674–83. doi: 10.1091/mbc.e05-07-0659
20. Suzuki M, Otsuka T, Ohsaki Y, Cheng J, Taniguchi T, Hashimoto H, et al. Derlin-1 and UBXD8 Are Engaged in Dislocation and Degradation of Lipidated ApoB-100 at Lipid Droplets. *Mol Biol Cell* (2012) 23(5):800–10. doi: 10.1091/mbc.e11-11-0950
21. Jarc E, Petan T. A Twist of FATe: Lipid Droplets and Inflammatory Lipid Mediators. *Biochimie* (2020) 169:69–87. doi: 10.1016/j.biochi.2019.11.016
22. Accioly MT, Pacheco P, Maya-Monteiro CM, Carrossini N, Robbs BK, Oliveira SS, et al. Lipid Bodies are Reservoirs of Cyclooxygenase-2 and Sites of prostaglandin-E2 Synthesis in Colon Cancer Cells. *Cancer Res* (2008) 68 (6):1732–40. doi: 10.1158/0008-5472.CAN-07-1999
23. Silva AR, Pacheco P, Vieira-de-Abreu A, Maya-Monteiro CM, D'Alegria B, Magalhaes KG, et al. Lipid Bodies in Oxidized LDL-induced Foam Cells Are Leukotriene-Synthesizing Organelles: A MCP-1/CCL2 Regulated Phenomenon. *Biochim Biophys Acta* (2009) 1791(11):1066–75. doi: 10.1016/j.bbalip.2009.06.004
24. Brasaemle DL, Wolins NE. Isolation of Lipid Droplets From Cells by Density Gradient Centrifugation. *Curr Protoc Cell Biol* (2016) 72:3.15.1–3.3. doi: 10.1002/cpcb.10
25. Harris LA, Shew TM, Skinner JR, Wolins NE. A Single Centrifugation Method for Isolating Fat Droplets From Cells and Tissues. *J Lipid Res* (2012) 53(5):1021–5. doi: 10.1194/jlr.D023598
26. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat Methods* (2012) 9(7):671–5. doi: 10.1038/nmeth.2089

27. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed Maxlfq. *Mol Cell Proteomics* (2014) 13(9):2513–26. doi: 10.1074/mcp.M113.031591
28. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated Into the MaxQuant Environment. *J Proteome Res* (2011) 10(4):1794–805. doi: 10.1021/pr101065j
29. Cox J, Mann M. MaxQuant Enables High Peptide Identification Rates, Individualized P.P.B.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat Biotechnol* (2008) 26(12):1367–72. doi: 10.1038/nbt.1511
30. Huang da W, Sherman BT, Lempicki RA. Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat Protoc* (2009) 4(1):44–57. doi: 10.1038/nprot.2008.211
31. Huang da W, Sherman BT, Lempicki RA. Bioinformatics Enrichment Tools: Paths Toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res* (2009) 37(1):1–13. doi: 10.1093/nar/gkn923
32. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. String v11: Protein-Protein Association Networks With Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res* (2019) 47(D1):D607–d13. doi: 10.1093/nar/gky1131
33. Kimmel AR, Sztalryd C. The Perilipins: Major Cytosolic Lipid Droplet-Associated Proteins and Their Roles in Cellular Lipid Storage, Mobilization, and Systemic Homeostasis. *Annu Rev Nutr* (2016) 36:471–509. doi: 10.1146/annurev-nutr-071813-105410
34. Lee B, Zhu J, Wolins NE, Cheng JX, Buhman KK. Differential Association of Adipophilin and TIP47 Proteins With Cytoplasmic Lipid Droplets in Mouse Enterocytes During Dietary Fat Absorption. *Biochim Biophys Acta* (2009) 1791(12):1173–80. doi: 10.1016/j.bbalip.2009.08.002
35. Zhang S, Wang Y, Cui L, Deng Y, Xu S, Yu J, et al. Morphologically and Functionally Distinct Lipid Droplet Subpopulations. *Sci Rep* (2016) 6:29539. doi: 10.1038/srep29539
36. Fujimoto Y, Itabe H, Sakai J, Makita M, Noda J, Mori M, et al. Identification of Major Proteins in the Lipid Droplet-Enriched Fraction Isolated From the Human Hepatocyte Cell Line Huh7. *Biochim Biophys Acta* (2004) 1644 (1):47–59. doi: 10.1016/j.bbamcr.2003.10.018
37. Alemayehu M, Dragan M, Pape C, Siddiqui I, Sacks DB, Di Guglielmo GM, et al. beta-Arrestin2 Regulates Lysophosphatidic Acid-Induced Human Breast Tumor Cell Migration and Invasion *Via* Rap1 and IQGAP1. *PloS One* (2013) 8(2):e56174. doi: 10.1371/journal.pone.0056174
38. White CD, Li Z, Dillon DA, Sacks DB. IQGAP1 Protein Binds Human Epidermal Growth Factor Receptor 2 (HER2) and Modulates Trastuzumab Resistance. *J Biol Chem* (2011) 286(34):29734–47. doi: 10.1074/jbc.M111.220939
39. Jadeski L, Mataraza JM, Jeong HW, Li Z, Sacks DB. IQGAP1 Stimulates Proliferation and Enhances Tumorigenesis of Human Breast Epithelial Cells. *J Biol Chem* (2008) 283(2):1008–17. doi: 10.1074/jbc.M708466200
40. Sun G, Liu Y, Wang K, Xu Z. miR-506 Regulates Breast Cancer Cell Metastasis by Targeting IQGAP1. *Int J Oncol* (2015) 47(5):1963–70. doi: 10.3892/ijo.2015.3161
41. Casteel DE, Turner S, Schwappacher R, Rangaswami H, Su-Yuo J, Zhuang S, et al. Rho Isoform-Specific Interaction With IQGAP1 Promotes Breast Cancer Cell Proliferation and Migration. *J Biol Chem* (2012) 287 (45):38367–78. doi: 10.1074/jbc.M112.377499
42. Cho CY, Lee KT, Chen WC, Wang CY, Chang YS, Huang HL, et al. MST3 Promotes Proliferation and Tumorigenicity Through the VAV2/Rac1 Signal Axis in Breast Cancer. *Oncotarget* (2016) 7(12):14586–604. doi: 10.18632/oncotarget.7542
43. Clarke CJ, Gross SR, Ismail TM, Rudland PS, Al-Medhtiy M, Santangeli M, et al. Activation of Tissue Plasminogen Activator by Metastasis-Inducing S100P Protein. *Biochem J* (2017) 474(19):3227–40. doi: 10.1042/BCJ20170578
44. Zhou C, Zhong Q, Rhodes LV, Townley I, Bratton MR, Zhang Q, et al. Proteomic Analysis of Acquired Tamoxifen Resistance in MCF-7 Cells Reveals Expression Signatures Associated With Enhanced Migration. *Breast Cancer Res* (2012) 14(2):R45. doi: 10.1186/bcr3144
45. Kikuchi K, McNamara KM, Miki Y, Iwabuchi E, Kanai A, Miyashita M, et al. S100P and Ezrin Promote Trans-Endothelial Migration of Triple Negative Breast Cancer Cells. *Cell Oncol (Dordr)* (2019) 42(1):67–80. doi: 10.1007/s13402-018-0408-2
46. Wang CQ, Li Y, Huang BF, Zhao YM, Yuan H, Guo D, et al. EGFR Conjunct FSCN1 as a Novel Therapeutic Strategy in Triple-Negative Breast Cancer. *Sci Rep* (2017) 7(1):15654. doi: 10.1038/s41598-017-15939-9
47. Wang CQ, Tang CH, Wang Y, Jin L, Wang Q, Li X, et al. FSCN1 Gene Polymorphisms: Biomarkers for the Development and Progression of Breast Cancer. *Sci Rep* (2017) 7(1):15887. doi: 10.1038/s41598-017-16196-6
48. Acharya S, Yao J, Li P, Zhang C, Lowery FJ, Zhang Q, et al. Sphingosine Kinase 1 Signaling Promotes Metastasis of Triple-Negative Breast Cancer. *Cancer Res* (2019) 79(16):4211–26. doi: 10.1158/0008-5472.CAN-18-3803
49. Snyder M, Huang XY, Zhang JJ. Signal Transducers and Activators of Transcription 3 (STAT3) Directly Regulates Cytokine-Induced Fascin Expression and Is Required for Breast Cancer Cell Migration. *J Biol Chem* (2011) 286(45):38886–93. doi: 10.1074/jbc.M111.286245
50. Al-Alwan M, Olabi S, Ghebeh H, Barhoush E, Tulbah A, Al-Tweigeri T, et al. Fascin Is a Key Regulator of Breast Cancer Invasion That Acts *Via* the Modification of Metastasis-Associated Molecules. *PloS One* (2011) 6(11): e27339. doi: 10.1371/journal.pone.0027339
51. Zhao H, Kang X, Xia X, Wo L, Gu X, Hu Y, et al. miR-145 Suppresses Breast Cancer Cell Migration by Targeting FSCN-1 and Inhibiting Epithelial-Mesenchymal Transition. *Am J Transl Res* (2016) 8(7):3106–14. doi: 1943-8141/AJTR0019453
52. Wu D, Haruta A, Wei Q. GIPC1 Interacts With MyoGEF and Promotes MDA-MB-231 Breast Cancer Cell Invasion. *J Biol Chem* (2010) 285 (37):28643–50. doi: 10.1074/jbc.M110.107649
53. Chittenden TW, Pak J, Rubio R, Cheng H, Holton K, Prendergast N, et al. Therapeutic Implications of GIPC1 Silencing in Cancer. *PloS One* (2010) 5 (12):e15581. doi: 10.1371/journal.pone.0015581
54. Jiang C, Veon W, Li H, Hallows KR, Roy P. Epithelial Morphological Reversion Drives Profilin-1-induced Elevation of p27(kip1) in Mesenchymal Triple-Negative Human Breast Cancer Cells Through AMP-Activated Protein Kinase Activation. *Cell Cycle* (2015) 14(18):2914–23. doi: 10.1080/15384101.2015.1069929
55. Chakraborty S, Jiang C, Gau D, Oddo M, Ding Z, Vollmer L, et al. Profilin-1 Deficiency Leads to SMAD3 Upregulation and Impaired 3D Outgrowth of Breast Cancer Cells. *Br J Cancer* (2018) 119(9):1106–17. doi: 10.1038/s41416-018-0284-6
56. Ding Z, Joy M, Bhargava R, Gunsaulus M, Lakshman N, Miron-Mendoza M, et al. Profilin-1 Downregulation has Contrasting Effects on Early vs Late Steps of Breast Cancer Metastasis. *Oncogene* (2014) 33(16):2065–74. doi: 10.1038/onc.2013.166
57. Zou L, Jaramillo M, Whaley D, Wells A, Panchapakesa V, Das T, et al. Profilin-1 Is a Negative Regulator of Mammary Carcinoma Aggressiveness. *Br J Cancer* (2007) 97(10):1361–71. doi: 10.1038/sj.bjc.6604038
58. Lin H, Zhang H, Wang J, Lu M, Zheng F, Wang C, et al. A Novel Human Fab Antibody for Trop2 Inhibits Breast Cancer Growth *In Vitro* and In Vivo. *Int J Cancer* (2014) 134(5):1239–49. doi: 10.1002/ijc.28451
59. Guerra E, Trerotola M, Tripaldi R, Aloisi AL, Simeone P, Sacchetti A, et al. Trop-2 Induces Tumor Growth Through AKT and Determines Sensitivity to AKT Inhibitors. *Clin Cancer Res* (2016) 22(16):4197–205. doi: 10.1158/1078-0432.CCR-15-1701
60. Pu T, Shen M, Li S, Yang L, Gao H, Xiao L, et al. Repression of miR-135b-5p Promotes Metastasis of Early-Stage Breast Cancer by Regulating Downstream Target SDCBP. *Lab Invest* (2019) 99(9):1296–308. doi: 10.1038/s41374-019-0258-1
61. Zhang J, Qian X, Liu F, Guo X, Gu F, Fu L. Silencing of Syndecan-Binding Protein Enhances the Inhibitory Effect of Tamoxifen and Increases Cellular Sensitivity to Estrogen. *Cancer Biol Med* (2018) 15(1):29–38. doi: 10.20892/j.issn.2095-3941.2017.0122
62. Qian XL, Li YQ, Yu B, Gu F, Liu FF, Li WD, et al. Syndecan Binding Protein (SDCBP) is Overexpressed in Estrogen Receptor Negative Breast Cancers, and Is a Potential Promoter for Tumor Proliferation. *PloS One* (2013) 8(3): e60046. doi: 10.1371/journal.pone.0060046
63. Liu J, Yang Y, Wang H, Wang B, Zhao K, Jiang W, et al. Syntenin1/MDA-9 (SDCBP) Induces Immune Evasion in Triple-Negative Breast Cancer by Upregulating PD-L1. *Breast Cancer Res Treat* (2018) 171(2):345–57. doi: 10.1007/s10549-018-4833-8

64. Xu H, Qian M, Zhao B, Wu C, Maskey N, Song H, et al. Inhibition of RAB1A Suppresses Epithelial-Mesenchymal Transition and Proliferation of Triple-Negative Breast Cancer Cells. *Oncol Rep* (2017) 37(3):1619–26. doi: 10.3892/or.2017.5404
65. Zhang W, Xu J, Wang K, Tang X, He J. miR1393p Suppresses the Invasion and Migration Properties of Breast Cancer Cells by Targeting RAB1A. *Oncol Rep* (2019) 42(5):1699–708. doi: 10.3892/or.2019.7297
66. Roovers K, Wagner S, Storbeck CJ, O'Reilly P, Lo V, Northey JJ, et al. The Ste20-like Kinase SLK Is Required for ErbB2-Driven Breast Cancer Cell Motility. *Oncogene* (2009) 28(31):2839–48. doi: 10.1038/onc.2009.146
67. Molinie N, Rubtsova SN, Fokin A, Visweshwaran SP, Rocques N, Polesskaya A, et al. Cortical Branched Actin Determines Cell Cycle Progression. *Cell Res* (2019) 29(6):432–45. doi: 10.1038/s41422-019-0160-9
68. Chang YW, Chen HA, Tseng CF, Hong CC, Ma JT, Hung MC, et al. De-Acetylation and Degradation of HSPA5 Is Critical for E1A Metastasis Suppression in Breast Cancer Cells. *Oncotarget* (2014) 5(21):10558–70. doi: 10.18632/oncotarget.2510
69. Chang YW, Tseng CF, Wang MY, Chang WC, Lee CC, Chen LT, et al. Deacetylation of HSPA5 by HDAC6 Leads to GP78-Mediated HSPA5 Ubiquitination at K447 and Suppresses Metastasis of Breast Cancer. *Oncogene* (2016) 35(12):1517–28. doi: 10.1038/onc.2015.214
70. Cook KL, Clarke R. Heat Shock 70 kDa Protein 5/Glucose-Regulated Protein 78 "Amp"Ing Up Autophagy. *Autophagy* (2012) 8(12):1827–9. doi: 10.4161/auto.21765
71. Dong X, Liu F, Sun L, Liu M, Li D, Su D, et al. Oncogenic Function of Microtubule End-Binding Protein 1 in Breast Cancer. *J Pathol* (2010) 220 (3):361–9. doi: 10.1002/path.2662
72. Hu X, Guo J, Zheng L, Li C, Zheng TM, Tanyi JL, et al. The Heterochronic microRNA Let-7 Inhibits Cell Motility by Regulating the Genes in the Actin Cytoskeleton Pathway in Breast Cancer. *Mol Cancer Res* (2013) 11(3):240–50. doi: 10.1158/1541-7786.MCR-12-0432
73. Asp N, Kvalvaag A, Sandvig K, Pust S. Regulation of ErbB2 Localization and Function in Breast Cancer Cells by ERM Proteins. *Oncotarget;* (2016) p:25443–60. doi: 10.18632/oncotarget.8327
74. Huang P, Liao R, Chen X, Wu X, Li X, Wang Y, et al. Nuclear Translocation of PLSCR1 Activates STAT1 Signaling in Basal-Like Breast Cancer. *Theranostics* (2020) 10(10):4644–58. doi: 10.7150/thno.43150
75. Varikuti S, Oghumu S, Elbaz M, Volpedo G, Ahirwar DK, Alarcon PC, et al. STAT1 Gene Deficient Mice Develop Accelerated Breast Cancer Growth and Metastasis Which Is Reduced by IL-17 Blockade. *Oncoimmunology* (2017) 6 (11):e1361088. doi: 10.1080/2162402X.2017.1361088
76. Chan SR, Vermi W, Luo J, Lucini L, Rickert C, Fowler AM, et al. STAT1-Deficient Mice Spontaneously Develop Estrogen Receptor Alpha-Positive Luminal Mammary Carcinomas. *Breast Cancer Res* (2012) 14(1):R16. doi: 10.1186/bcr3100
77. Hix LM, Karavitis J, Khan MW, Shi YH, Khazaie K, Zhang M. Tumor STAT1 Transcription Factor Activity Enhances Breast Tumor Growth and Immune Suppression Mediated by Myeloid-Derived Suppressor Cells. *J Biol Chem* (2013) 288(17):11676–88. doi: 10.1074/jbc.M112.441402
78. Giudetti AM, De Domenico S, Ragusa A, Lunetti P, Gaballo A, Franck J, et al. A Specific Lipid Metabolic Profile Is Associated With the Epithelial Mesenchymal Transition Program. *Biochim Biophys Acta Mol Cell Biol Lipids* (2019) 1864(3):344–57. doi: 10.1016/j.bbalip.2018.12.011
79. Wright HJ, Hou J, Xu B, Cortez M, Potma EO, Tromberg BJ, et al. CDCP1 Drives Triple-Negative Breast Cancer Metastasis Through Reduction of Lipid-Droplet Abundance and Stimulation of Fatty Acid Oxidation. *Proc Natl Acad Sci USA* (2017) 114(32):E6556–e65. doi: 10.1073/pnas.1703791114
80. Sevinsky CJ, Khan F, Kokabee L, Darehshouri A, Maddipati KR, Conklin DS. NDRG1 Regulates Neutral Lipid Metabolism in Breast Cancer Cells. *Breast Cancer Res* (2018) 20(1):55. doi: 10.1186/s13058-018-0980-4
81. Santner SJ, Dawson PJ, Tait L, Soule HD, Eliason J, Mohamed AN, et al. Malignant MCF10CA1 Cell Lines Derived From Premalignant Human Breast Epithelial MCF10AT Cells. *Breast Cancer Res Treat* (2001) 65 (2):101–10. doi: 10.1023/A:1006461422273
82. Wilmanski T, Buhman K, Donkin SS, Burgess JR, Teegarden D. 1alpha,25-Dihydroxyvitamin D Inhibits *De Novo* Fatty Acid Synthesis and Lipid Accumulation in Metastatic Breast Cancer Cells Through Down-Regulation of Pyruvate Carboxylase. *J Nutr Biochem* (2017) 40:194–200. doi: 10.1016/j.jnutbio.2016.11.006
83. Suzuki M, Shinohara Y, Ohsaki Y, Fujimoto T. Lipid Droplets: Size Matters. *J Electron Microsc (Tokyo)* (2011) 60 Suppl 1:S101–16. doi: 10.1093/jmicro/dfr016
84. D'Aquila T, Zembroski AS, Buhman KK. Diet Induced Obesity Alters Intestinal Cytoplasmic Lipid Droplet Morphology and Proteome in the Postprandial Response to Dietary Fat. *Front Physiol* (2019) 10:180. doi: 10.3389/fphys.2019.00180
85. Daemen S, van Polanen N, Hesselink MKC. The Effect of Diet and Exercise on Lipid Droplet Dynamics in Human Muscle Tissue. *J Exp Biol* (2018) 221 (Pt Suppl 1):e167015. doi: 10.1242/jeb.167015
86. Schott MB, Weller SG, Schulze RJ, Krueger EW, Drizyte-Miller K, Casey CA, et al. Lipid Droplet Size Directs Lipolysis and Lipophagy Catabolism in Hepatocytes. *J Cell Biol* (2019) 218(10):3320–35. doi: 10.1083/jcb.201803153
87. Ding Y, Zhang S, Yang L, Na H, Zhang P, Zhang H, et al. Isolating Lipid Droplets From Multiple Species. *Nat Protoc* (2013) 8(1):43–51. doi: 10.1038/nprot.2012.142
88. Thiam AR, Beller M. The Why, When and How of Lipid Droplet Diversity. *J Cell Sci* (2017) 130(2):315–24. doi: 10.1242/jcs.192021
89. Benador IY, Veliova M, Mahdaviani K, Petcherski A, Wikstrom JD, Assali EA, et al. Mitochondria Bound to Lipid Droplets Have Unique Bioenergetics, Composition, and Dynamics That Support Lipid Droplet Expansion. *Cell Metab* (2018) 27(4):869–85.e6. doi: 10.1016/j.cmet.2018.03.003
90. Sztalryd C, Brasaemle DL. The Perilipin Family of Lipid Droplet Proteins: Gatekeepers of Intracellular Lipolysis. *Biochim Biophys Acta Mol Cell Biol Lipids* (2017) 1862(10 Pt B):1221–32. doi: 10.1016/j.bbalip.2017.07.009
91. Wilfling F, Wang H, Haas JT, Krahmer N, Gould TJ, Uchida A, et al. Triacylglycerol Synthesis Enzymes Mediate Lipid Droplet Growth by Relocalizing From the ER to Lipid Droplets. *Dev Cell* (2013) 24(4):384–99. doi: 10.1016/j.devcel.2013.01.013
92. Krahmer N, Guo Y, Wilfling F, Hilger M, Lingrell S, Heger K, et al. Phosphatidylcholine Synthesis for Lipid Droplet Expansion Is Mediated by Localized Activation of CTP:phosphocholine Cytidylyltransferase. *Cell Metab* (2011) 14(4):504–15. doi: 10.1016/j.cmet.2011.07.013
93. Guo Y, Walther TC, Rao M, Stuurman N, Goshima G, Terayama K, et al. Functional Genomic Screen Reveals Genes Involved in Lipid-Droplet Formation and Utilization. *Nature* (2008) 453(7195):657–61. doi: 10.1038/nature06928
94. Smirnova E, Goldberg EB, Makarova KS, Lin L, Brown WJ, Jackson CL. ATGL has a Key Role in Lipid Droplet/Adiposome Degradation in Mammalian Cells. *EMBO Rep* (2006) 7(1):106–13. doi: 10.1038/sj.embor.7400559
95. Ohashi M, Mizushima N, Kabeya Y, Yoshimori T. Localization of Mammalian NAD(P)H Steroid Dehydrogenase-Like Protein on Lipid Droplets. *J Biol Chem* (2003) 278(38):36819–29. doi: 10.1074/jbc.M301408200
96. Caldas H, Herman GE. NSDHL, an Enzyme Involved in Cholesterol Biosynthesis, Traffics Through the Golgi and Accumulates on ER Membranes and on the Surface of Lipid Droplets. *Hum Mol Genet* (2003) 12(22):2981–91. doi: 10.1093/hmg/ddg321
97. Antalis CJ, Arnold T, Rasool T, Lee B, Buhman KK, Siddiqui RA. High ACAT1 Expression in Estrogen Receptor Negative Basal-Like Breast Cancer Cells is Associated With LDL-Induced Proliferation. *Breast Cancer Res Treat* (2010) 122(3):661–70. doi: 10.1007/s10549-009-0594-8
98. Antalis CJ, Uchida A, Buhman KK, Siddiqui RA. Migration of MDA-MB-231 Breast Cancer Cells Depends on the Availability of Exogenous Lipids and Cholesterol Esterification. *Clin Exp Metastasis* (2011) 28(8):733–41. doi: 10.1007/s10585-011-9405-9
99. Chimento A, Casaburi I, Avena P, Trotta F, De Luca A, Rago V, et al. Cholesterol and Its Metabolites in Tumor Growth: Therapeutic Potential of Statins in Cancer Treatment. *Front Endocrinol (Lausanne)* (2018) 9:807. doi: 10.3389/fendo.2018.00807
100. Huang B, B-l S, Xu C. Cholesterol Metabolism in Cancer: Mechanisms and Therapeutic Opportunities. *Nat Metabol* (2020) p:132–41. doi: 10.1038/s42255-020-0174-0
101. Sharpe LJ, Brown AJ. Controlling Cholesterol Synthesis Beyond 3-Hydroxy-3-methylglutaryl-CoA Reductase (HMGCR). *J Biol Chem* (2013) 288 (26):18707–15. doi: 10.1074/jbc.R113.479808

102. Yoon SH, Kim HS, Kim RN, Jung SY, Hong BS, Kang EJ, et al. NAD(P)-Dependent Steroid Dehydrogenase-Like Is Involved in Breast Cancer Cell Growth and Metastasis. *BMC Cancer* (2020) 20(1):375. doi: 10.1186/s12885-020-06840-2
103. Feltrin S, Ravera F, Traversone N, Ferrando L, Bedognetti D, Ballestrero A, et al. Sterol Synthesis Pathway Inhibition as a Target for Cancer Treatment. *Cancer Lett* (2020) 493:19–30. doi: 10.1016/j.canlet.2020.07.010
104. Leber R, Landl K, Zinser E, Ahorn H, Spök A, Kohlwein SD, et al. Dual Localization of Squalene Epoxidase, Erg1p, in Yeast Reflects a Relationship Between the Endoplasmic Reticulum and Lipid Particles. *Mol Biol Cell* (1998) 9(2):375–86. doi: 10.1091/mbc.9.2.375
105. Ta MT, Kapterian TS, Fei W, Du X, Brown AJ, Dawes IW, et al. Accumulation of Squalene Is Associated With the Clustering of Lipid Droplets. *FEBS J* (2012) 279(22):4231–44. doi: 10.1111/febs.12015
106. Polycarpou-Schwarz M, Groß M, Mestdagh P, Schott J, Grund SE, Hildenbrand C, et al. The Cancer-Associated Microprotein CASIMO1 Controls Cell Proliferation and Interacts With Squalene Epoxidase Modulating Lipid Droplet Formation. *Oncogene* (2018) 37(34):4750–68. doi: 10.1038/s41388-018-0281-5
107. Brown DN, Caffa I, Cirmena G, Piras D, Garuti A, Gallo M, et al. Squalene Epoxidase Is a Bona Fide Oncogene by Amplification With Clinical Relevance in Breast Cancer. *Sci Rep* (2016) 6:19435. doi: 10.1038/srep19435
108. Guo YJ, Pan WW, Liu SB, Shen ZF, Xu Y, Hu LL. ERK/MAPK Signalling Pathway and Tumorigenesis. *Exp Ther Med* (2020) 19(3):1997–2007. doi: 10.3892/etm.2020.8454
109. Sui Z, Zhou J, Cheng Z, Lu P. Squalene Epoxidase (SQLE) Promotes the Growth and Migration of the Hepatocellular Carcinoma Cells. *Tumour Biol* (2015) 36(8):6173–9. doi: 10.1007/s13277-015-3301-x
110. Ge H, Zhao Y, Shi X, Tan Z, Chi X, He M, et al. Squalene Epoxidase Promotes the Proliferation and Metastasis of Lung Squamous Cell Carcinoma Cells Though Extracellular Signal-Regulated Kinase Signaling. *Thorac Cancer* (2019) 10(3):428–36. doi: 10.1111/1759-7714.12944
111. Qin Y, Hou Y, Liu S, Zhu P, Wan X, Zhao M, et al. A Novel Long Non-Coding RNA Lnc030 Maintains Breast Cancer Stem Cell Stemness by Stabilizing Sqle mRNA and Increasing Cholesterol Synthesis. *Adv Sci (Weinh)* (2021) 8(2):2002232. doi: 10.1002/advs.202002232
112. Valastyan S, Weinberg RA. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* (2011) 147(2):275–92. doi: 10.1016/j.cell.2011.09.024
113. Wurth L. Versatility of RNA-Binding Proteins in Cancer. *Comp Funct Genomics* (2012) 2012:178525. doi: 10.1155/2012/178525
114. Mohibi S, Chen X, Zhang J. Cancer The'rbp'eutics-RNA-Binding Proteins as Therapeutic Targets for Cancer. *Pharmacol Ther* (2019) 203:107390. doi: 10.1016/j.pharmthera.2019.07.001
115. Gu W, Pan F, Singer RH. Blocking Beta-Catenin Binding to the ZBP1 Promoter Represses ZBP1 Expression, Leading to Increased Proliferation and Migration of Metastatic Breast-Cancer Cells. *J Cell Sci* (2009) 122(Pt 11):1895–905. doi: 10.1242/jcs.045278
116. Ebright RY, Lee S, Wittner BS, Niederhoffer KL, Nicholson BT, Bardia A, et al. Deregulation of Ribosomal Protein Expression and Translation Promotes Breast Cancer Metastasis. *Science* (2020) 367(6485):1468–73. doi: 10.1126/science.aay0939
117. Dvorak AM, Morgan ES, Weller PF. RNA Is Closely Associated With Human Mast Cell Lipid Bodies. *Histol Histopathol* (2003) 18(3):943–68. doi: 10.14670/HH-18.943
118. Wan HC, Melo RC, Jin Z, Dvorak AM, Weller PF. Roles and Origins of Leukocyte Lipid Bodies: Proteomic and Ultrastructural Studies. *FASEB J* (2007) 21(1):167–78. doi: 10.1096/fj.06-6711com
119. Schuldiner M, Bohnert M. A Different Kind of Love - Lipid Droplet Contact Sites. *Biochim Biophys Acta* (2017) 1862(10 Pt B):1188–96. doi: 10.1016/j.bbalip.2017.06.005

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 Zembroski, Andolino, Buhman and Teegarden. 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.2021.621810

Coupling of Cell Surface Biotinylation and SILAC-Based Quantitative Proteomics Identified Myoferlin as a Potential Therapeutic Target for Nasopharyngeal Carcinoma Metastasis

Maoyu Li[1,2†], Fang Peng[2†], Guoqiang Wang[2†], Xujun Liang[2], Meiying Shao[2], Zhuchu Chen[1,2] and Yongheng Chen[1,2,3*]*

[1] Department of Gastroenterology, Xiangya Hospital, Central South University, Changsha, China, [2] NHC Key Laboratory of Cancer Proteomics, Xiangya Hospital, Central South University, Changsha, China, [3] Key Laboratory of Medical Genetics and College of Life Sciences, Central South University, Changsha, China

OPEN ACCESS

Edited by:
Suman S. Thakur, Centre for Cellular and Molecular Biology (CCMB), India

Reviewed by:
Xue Xiao, Guangxi Medical University, China
Xiaoying Zhou, Guangxi Medical University, China

***Correspondence:**
Yongheng Chen
yonghenc@163.com
Zhuchu Chen
chenzhuchu@126.com

† These authors have contributed equally to this work

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

Received: *27 October 2020*
Accepted: *30 April 2021*
Published: *09 June 2021*

Citation:
Li M, Peng F, Wang G, Liang X, Shao M, Chen Z and Chen Y (2021) Coupling of Cell Surface Biotinylation and SILAC-Based Quantitative Proteomics Identified Myoferlin as a Potential Therapeutic Target for Nasopharyngeal Carcinoma Metastasis. Front. Cell Dev. Biol. 9:621810. doi: 10.3389/fcell.2021.621810

Distant metastasis is a major cause of treatment failure in nasopharyngeal carcinoma (NPC) patients. Cell surface proteins represent attractive targets for cancer diagnosis or therapy. However, the cell surface proteins associated with NPC metastasis are poorly understood. To identify potential therapeutic targets for NPC metastasis, we isolated cell surface proteins from two isogenic NPC cell lines, 6-10B (low metastatic) and 5-8F (highly metastatic), through cell surface biotinylation. Stable isotope labeling by amino acids in cell culture (SILAC) based proteomics was applied to comprehensively characterize the cell surface proteins related with the metastatic phenotype. We identified 294 differentially expressed cell surface proteins, including the most upregulated protein myoferlin (MYOF), two receptor tyrosine kinases(RTKs) epidermal growth factor receptor (EGFR) and ephrin type-A receptor 2 (EPHA2) and several integrin family molecules. These differentially expressed proteins are enriched in multiple biological pathways such as the FAK-PI3K-mTOR pathway, focal adhesions, and integrin-mediated cell adhesion. The knockdown of MYOF effectively suppresses the proliferation, migration and invasion of NPC cells. Immunohistochemistry analysis also showed that MYOF is associated with NPC metastasis. We experimentally confirmed, for the first time, that MYOF can interact with EGFR and EPHA2. Moreover, MYOF knockdown could influence not only EGFR activity and its downstream epithelial–mesenchymal transition (EMT), but also EPHA2 ligand-independent activity. These findings suggest that MYOF might be an attractive potential therapeutic target that has double effects of simultaneously influencing EGFR and EPHA2 signaling pathway. In conclusion, this is the first study to profile the cell surface proteins associated with NPC metastasis and provide valuable resource for future researches.

Keywords: nasopharyngeal carcinoma, cell surface proteins, metastasis, biotinylation, SILAC

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is an aggressive malignancy common in Southern China and Southeast Asia. NPC is prone to early metastasis, but patients are asymptomatic at the early stage (Chua et al., 2016). Therefore, most NPC patients (60–70% cases) are diagnosed at an advanced stage (Mao et al., 2009). At present, radiotherapy and chemo-radiotherapy are the routine treatment strategies for patients with non-metastatic or locally advanced NPC. It has been reported that for approximately 25-30% of patients therapy still fails, with local recurrence and/or distant metastases, and the prognosis for patients with metastatic disease is poor (Lai et al., 2011). NPC metastasis and recurrence are the main bottleneck for treatment. Therefore, the demand for new therapeutic strategies for advanced NPC, to further improve treatment outcome is urgent.

The interaction between cancer cells and surrounding stromal cells in the tumor microenvironment plays a key role in cancer cell migration, invasion and metastasis. Cell surface proteins are important player in these processes. Cell surface proteins, primarily represented by plasma membrane proteins, account for approximately one-third of the proteins encoded by the human genome (Vuckovic et al., 2013). These proteins have a variety of important biological functions, such as signal transduction, cell-to-cell recognition, and material transport. Given their accessibility and significant biological function, they have become ideal targets for novel therapies. At present, more than two-thirds of the current protein-based drug targets are plasma membrane proteins (Overington et al., 2006).

Due to their low abundance and poor solubility, routine proteomics analysis strategies are insufficient for identification and characterization of these proteins. Therefore, strategies combining cell surface enrichment and quantitative proteomics, which can reduce sample complexity, resulting in identification of more intermediate to low abundant proteins, have been widely exploited. Several methods, such as ultracentrifugation, cell-surface biotinylation and cell surface capture, have been used to extract or enrich the cell surface and/or plasma membrane proteins. Cell-surface biotinylation is one of the most commonly used methods for isolation of plasma membrane proteins, in which the extracellular domains of integral and plasma membrane-associated proteins are covalently labeled with a reactive biotin ester. Through tagging with a membrane-impermeable biotin, cell surface proteins can be separated from cell extracts with streptavidin-linked beads. This approach has been widely used to separate plasma membrane proteins associated with tumor metastasis (Lund et al., 2009), tumor cell cycle (Özlü et al., 2015), and T cell-mediated autoimmunity (Buehler et al., 2018), which demonstrates that it can improve throughput and selectivity for detection of membrane proteins in low abundance.

However, the cell surface proteins involved in NPC metastasis have not been well investigated. Here, Stable isotope labeling by amino acids in cell culture (SILAC) was employed to compare cell surface proteins enriched via surface biotinylation from a pair of isogenic NPC cells with low and high metastatic potential. A total of 1029 proteins localized at the cell surface were identified. Of the quantified cell surface proteins, 294 differentially expressed proteins were enriched in multiple pathways such as the phosphatidylinositol 3-kinase - protein kinase B (PI3K-AKT) signaling pathway, integrin pathway. We verified that the most upregulated membrane protein myoferlin (MYOF) was metastasis-related protein whose knockdown suppressed the malignant phenotype of NPC tumor cells. We further, for the first time experimentally confirmed that the interactions of MYOF with epidermal growth factor receptor (EGFR) and ephrin type-A receptor 2 (EPHA2), both of which are metastasis-related receptor tyrosine kinases (RTK). We found that MYOF impairs not only EGFR activation and its downstream epithelial–mesenchymal transition (EMT) process, but also EPHA2 ligand-independent activation. Therefore, MYOF targeted therapy might have double effect that simultaneously influence EGFR and EPHA2 pathway in NPC. Collectively, this is the first large-scale characterization of NPC metastasis-related cell surface membrane proteins, which provide valuable resource for future researches.

MATERIALS AND METHODS

Cell Culture and Small Hairpin RNA (shRNA) Lentiviral Transduction

A isogenic pair of human malignant NPC cell lines, 5-8F (high metastatic potential) and 6-10B (low metastatic potential), were cultured in Roswell Park Memorial Institute 1640 (RPMI-1640) medium (Gibco BRL, Grand Island, NY, United States) supplemented with 10% fetal bovine serum (FBS) (Gibco) at 37°C in 5% CO_2.

For lentiviral transduction, cells were trypsinized and resuspended in RPMI-1640 containing 10% FBS at a density of 1×10^6 cells/ml. pLV-EGFP-T2A-Neo-U6-MYOF-shRNA or pLV-EGFP-T2A-Neo-U6-scramble-shRNA (Cyagen) was added to the trypsinized cell suspension at a multiplicity of infection (MOI) of 30. Cells were subjected to puromycin selection to generate stable MYOF-knockdown cell lines.

SILAC Labeling

For SILAC experiments, RPMI-1640 medium lacking L-lysine, [$^{13}C_6$,$^{15}N_2$]-L lysine/[$^{13}C_6$,$^{15}N_4$]-L Arginine (heavy label) and unlabeled L-lysine/L-Arginine (light label) were purchased form Thermo Scientific (Baltimore, MD, United States). Light and heavy RPMI-1640 media were prepared according to the standard RPMI-1640 formulation except with the supplement of dialyzed FBS (Invitrogen). The low metastatic cell line were grown in light medium and the high metastatic 5-8F cells were cultured in heavy medium for at least 6 divisions to allow complete incorporation of the isotope-labeled amino acids. These experiments were performed in triplicate.

Affinity Purification of Cell-Surface Membrane Proteins

Cell surface protein isolation was performed according to the manufacturer's instructions. Briefly, after cells attached to the

flasks, they were washed three times with PBS (pH 7.4) at 37°C to remove culture medium and FBS. A 1-mL solution of 0.5 mg/mL Sulfo-NHS-SS-biotin (Pierce, Rockford, IL, United States) in PBS was added to each flask, and the cells were incubated at 37°C for 10 min. Then, 1 mL of quenching solution was added to each flask and incubated with cells at room temperature for 5 min to quench the reaction, and then, the cells were washed with ice-cold PBS three times to remove excess reagent and byproducts. Biotinylated 6-10B and 5-8F cells were harvested into ice-cold PBS and then pooled in a 1:1 ratio. The mixed cells were lysed on ice for 30 min with gentle shaking. Cell lysates were centrifuged at 15000 × *g* for 30 min at 4°C, and the supernatants were collected. Avidin-agarose resin was added into spin columns and washed with washing buffer three times. The clarified cell lysates were placed into the spin columns. After a 60-min incubation at room temperature with end-over-end mixing, the resin was washed three times. The bound proteins were eluted with 50 mM DTT in SDS loading buffer. The elution process was repeated, and the eluates were combined and stored at −20°C for further analysis.

Enzymatic Digestion of Proteins

The eluted proteins were reduced with dithiothreitol (DTT) at a final concentration of 10 mM for 30 min at room temperature and then alkylated with iodoacetamide (IAA) at a final concentration of 50 mM for 30 min in the dark. The resulting solutions were diluted 10 times with 100 mM NH_4HCO_3 (pH 8.5) and digested with mass spectrometry (MS)-grade trypsin (Promega, Madison, WI, United States) at an enzyme-to-substrate ratio of 1/75 (w/w) at 37°C overnight.

MS and Data Acquisition

Peptide samples were analyzed in duplicate (two technical replicates). For each technical replicates, the peptide mixtures were fractionated into 5 or 10 fractions via strong anion exchange chromatography according to published procedures (Li et al., 2016). Each fraction was independently analyzed using an EASY-nano LC system (Proxeon Biosystems, Odense, Denmark) coupled online with an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA, United States). Briefly, peptides were loaded onto a PepMap C18 trap column (75 μm, 15 cm; Dionex Corp.) and eluted using a gradient from 100% solvent A (0.1% formic acid) to 35% solvent B (0.1% formic acid, 100% acetonitrile) for 38 min, 35 to 90% solvent B for 15 min, and 100% solvent B for 5 min (a total of 65 min at 200 nL/min). After each run, the column was washed with 90% solvent B and re-equilibrated with solvent A. Mass spectra were acquired in positive ion mode applying data-dependent automatic survey MS scan and tandem mass spectra (MS/MS) acquisition modes. Each MS scan in the Orbitrap analyzer (mass range = m/z 350–1800, resolution = 100,000) was followed by MS/MS of the seven most intense ions in the LTQ. Fragmentation in the LTQ was performed via high-energy, collision-activated dissociation, and selected sequenced ions were dynamically excluded for 30 s.

MS Data Processing and Quantitative Analysis

Raw MS files were processed using MaxQuant (version 1.5.8.30) (Cox and Mann, 2008). Carbamidomethylation (C) was set as a fixed modification, and oxidation (M), deamidation (N), N-acetyl (N-term), and DSP (K, N-term) were used as variable modifications. Missed tryptic cleavage was set to 2, and the minimal length required for a peptide was 6 amino acids. The tolerance of the precursor mass was 10 ppm, and the fragment mass tolerance was set to 0.5 Da. The false discovery rates (FDRs) for peptide and protein identification were both set to 0.01. All other parameters were set to the default settings. The datasets were searched against the UniProt human database (version 2019.07.06). Labeling was set to doublets of K0R0 and K8R10. For protein quantification, razor and unique peptides were used with two or more ratio counts. The protein SILAC ratio was calculated as the median of all SILAC peptide ratios.

Further analyses were performed using Perseus software (Tyanova et al., 2016). At least one unique peptide in at least one of the two technical replicates for each biological replicate was required for successful protein identification. Proteins that were identified in at least two biological replicates were selected for quantification analysis. Protein quantification was calculated by averaging the SILAC ratios of all replicate experiments. One-sample *t*-test was performed to evaluate the significance of the ratio. For calculation of correlation coefficients, datasets were filtered for entries with valid quantifications in all experiments. Differentially expressed protein groups were defined as those that had expression ratios larger than twofold and a p-Value < 0.05. Plasma membrane, extracellular, and cell surface proteins were annotated based on the GO (Gene Ontology) category Cellular compartment (CC) annotation using Perseus software.

Cell Proliferation Assay

The 5-8F cells were plated at 2×10^3 cells per well in 96-well tissue culture plates and cultured in 10% FBS complete RPMI-1640 medium for 6 days. Every 24 h, 20 μl of CCK8 reagent (5 mg/ml; Beyotime) was added to the wells, and cells were further incubated for 2 h. The absorbance of each well was read with a Bio-Tek Instruments EL310 Microplate Autoreader at 450 nm. Experiments were performed in triplicate.

Wound Healing Assay

Cell migration was determined with a scratch wound healing assay. Briefly, cells were grown to confluence in RPMI-1640 medium containing 10% FBS overnight in a 6-well plate. Cell monolayers were wounded by dragging a pipette tip through them. Cells were washed to remove cellular debris and allowed to migrate for 24-48 h. Images were taken at 0, 24, and 48 h after wounding under an inverted microscope.

Cell Migration Assay

Migration activity was measured with a Transwell assay (Corning, 3422). Approximately 5×10^4 cells were added to the upper chamber in 200 μl of 1% FBS RPMI-1640 medium.

The lower chamber contained 500 μl of 10% FBS complete RPMI-1640 medium. The plates were incubated for 24 h at 37°C in 5% CO_2. Cells were fixed in 3.7% formaldehyde solution for 15 min and stained with 0.05% crystal violet in PBS for 15 min. Cells on the upper side of the filters were removed with cotton-tipped swabs. Cells on the underside of the filters were examined and counted under a microscope. Each clone was plated in triplicate for each experiment, and each experiment was repeated three times.

Cell Invasion Assay

Invasion assays were performed in 24-well, 8-mm pore size Transwell chambers precoated with Matrigel (Corning) according to the manufacturer's instructions. The upper chamber was filled with 1×10^5 cells in RPMI-1640 medium containing 0.5% FBS. The lower chamber was filled with RPMI-1640 medium containing 10% FBS as a chemoattractant. The plates were incubated for 48 h at 37°C in 5% CO_2. Cells were fixed in 3.7% formaldehyde solution for 15 min and stained with 0.05% crystal violet in PBS for 15 min. Cells on the upper side of the filters were removed with cotton-tipped swabs. Cells on the underside of the filters were examined and counted under a microscope. Each clone was plated in triplicate for each experiment, and each experiment was repeated at least three times.

EGF-Mediated EMT Assay

The EGF-inducible EMT assay was performed according to previously described methods with minor modifications (Yip and Seow, 2012). Briefly, shMYOF-transfected 5-8F cells were trypsinized and plated at a density of 200,000/well in a 6-well plate. The culture medium was supplemented with EGF at a final concentration of 20 ng/mL. Cells were cultured further for 24 to 48 h. The expression levels of EMT markers were detected via western blotting.

Antibodies and Western Blotting

For immunostaining or western blotting, primary antibodies against the following proteins were used: MYOF (Santacruz); EGFR, pEGFR, EPHA2, and EPHA2-pS897 (Cell Signaling); VIM and E-cadherin (Abcam); and actin (Pierce). For western blot analysis, cells were washed once with ice-cold PBS and then lysed in cold RIPA buffer supplemented with protease inhibitor cocktail (Roche Applied Biosciences). Cell lysate was centrifuged at 12,000 × *g* for 30 min, boiled in 2 × loading sample buffer, separated on a 10% gradient SDS–PAGE gel and blotted onto a PVDF membrane. The membrane was blocked with 0.1% Tween-20 in TBS containing 5% non-fat milk and probed with the indicated primary antibody, followed by incubation with the appropriate secondary antibody conjugated to horseradish peroxidase (Sigma), and the signals were visualized via ECL (Millipore).

Co-immunoprecipitation (Co-IP)

The Co-IP experiment was performed as previously described (Sun et al., 2007). Briefly, cells were lysed in lysis buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, 1% NP-40, 1 mM Na_3VO_4, 1 mM PMSF) for 60 min on ice, followed by centrifuging at 12000 × *g* for 30 min at 4°C to remove cell debris. The whole cell lysates were precleared by incubation with Protein A/G Plus-Agarose (Santa Cruz Biotechnology) for 1 h at 4°C. The clarified supernatants were incubated with the indicated antibody or control (non-immune) serum overnight at 4°C, followed by incubation with Protein A/G Plus-Agarose (Santa Cruz Biotechnology) at 4°C for another 2 h. The agarose was washed 6 times and boiled in 2 × sodium dodecyl sulfate sample buffer for 5 min. The immunoprecipitated complexes were resolved using sodium dodecyl sulfate polyacrylamide gel electrophoresis and immunoblot with specific antibodies.

Total Cellular Protein and Membrane Fractionation

Cells were harvested in homogenization buffer A [20 mM TBS with 1 mM PMSF and complete EDTA-free protease inhibitor mixture (Roche)]. The cell suspension was homogenized 10 times on ice and then briefly centrifuged at 2,000 × *g* for 5 min at 4°C. The clarified lysate was then centrifuged at 80,000 rpm using a Beckman Ultracentrifuge (Beckman Coulter) for 1 h at 4°C. The membrane pellet was washed twice with buffer A and then centrifuged at 80,000 × *g* for 30 min at 4°C. The membrane pellet was dissolved in RIPA and stored in a refrigerator for further analysis.

Immunohistochemistry (IHC)

Fifty paraffin-embedded NPC tumor tissue specimens were collected from the archives in the Pathology Department (Xiangya Hospital, Central South University). IHC was performed according to the procedure described in our previous study (Xiao et al., 2010). Briefly, the sections were incubated with primary antibody overnight at 4°C and then with biotinylated secondary antibody, followed by addition of avidin-biotin peroxidase. Diaminobenzidine was used as the chromogen. The immunostaining intensity was evaluated as previously described with minor modification (Li et al., 2016). Briefly, each specimen was scored according to the intensity of staining and the area of staining. The intensity of staining was graded as follows: 0, no staining; 1 +, mild staining; 2 +, moderate staining; 3 +, intense staining. The area of staining was scored as follows: 0, no staining of cells in any microscopic field; 1 +, <30% of tumor cells stained positively; 2 +, between 30 and 60% stained positively; 3 +, >60% stained positively. A combined staining score (intensity + area, ranging from 0 to 6) was obtained for each case. A combined staining score of ≤ 4 was considered weak staining (low expression); a score > 4 was considered strong staining (high expression).

Bioinformatic Analysis

Differentially expressed proteins (DEPs) were subjected to Gene set enrichment analysis (GSEA) using GSEAPreranked tools in GSEA software (version 3.0) to gain insight into

over-represented Wikipathways[1]. The significance of enrichment was set to $p < 0.05$. Protein-protein interactions and their first neighbors for DEPs were retrieved using Harmonizome[2] to query Pathway Commons Protein-Protein Interactions dataset and imported into Cytoscape for visualization or presentation (Rouillard et al., 2016). Pathvisio (version 3.3[3]) was used to visualize the ratio values onto biological pathways obtained from Wikipathway. Student's *t*-test was used to analyze the significance of differences between groups. Correlations between MYOF expression and clinical characteristics associated with NPC tissues were analyzed with a Fisher's exact test. A *p*-Value < 0.05 was considered statistically significant. The Cancer Cell Line Encyclopedia(CCLE) proteomic data download from the website[4] (Nusinow et al., 2020).

RESULTS

Proteomics Analysis of Low and High Metastatic NPC Cell Surface Proteins

To minimize identification of non-surface-exposed proteins that had been labeled with Sulfo-NHS-SS-Biotin, we performed 3 biological replicate experiments for labeling and enrichment of cell surface proteins to examine the repeatedly enriched proteins which were more likely to be surface-exposed proteins (**Figure 1A**). In total, six experimental datasets were acquired. We calculated all pairwise Pearson correlations between these datasets. As shown in **Figure 1B**, the minimum Pearson correlation coefficient was 0.915, which demonstrates that the reproducibility of protein quantification was high at both the technical and biological level.

To increase the confidence of surface protein identification, we overlapped the three biological experimental datasets, and the proteins that were identified in at least two biological replicates (2161 proteins) were selected for subsequent analysis (**Figure 1C**). Of the 2161 proteins, 1029 proteins (47.6%) were annotated as plasma membrane, cell surface or extracellular region. As shown in **Figure 1D**, more than half (53%) of these proteins belonged in the PM or cell surface, which included transmembrane proteins such as integrins, cell adhesive molecules, receptors, and ion channels (**Supplementary Table 1**). The remaining (47%) proteins were membrane-associated proteins or interacted with the cell membrane according to GO annotation.

Analysis and Validation of Differentially Expressed Proteins

To identify and quantify differentially expressed proteins between the two cell lines, we performed three biological replicate experiments. A one sample *t*-test was used to evaluate the significance of differences. Of the 1029 proteins, 294 proteins were significantly differentially expressed (fold change > 2, *p*-Value < 0.05, **Figure 2A**, **Supplementary Table 2**). Among these differentially expressed proteins, 206 proteins (93 plasma membrane proteins, 41 cell surface proteins and 72 extracellular proteins) were upregulated in the high metastatic NPC cell line (**Table 1**). The remaining 88 proteins (48 plasma membrane proteins, 5 cell surface proteins and 35 extracellular proteins) were downregulated in the high metastatic cell line. **Table 2** shows the top 20 up- and downregulated proteins and their important information for identification, such as score, identified peptides, ratio, *p*-Value and subcellular location.

In order to validate the proteomic results, three differentially expressed cell surface proteins, the most upregulated protein MYOF, and two other typical plasma proteins, EGFR and EPHA2, were selected for western blotting and immunohistochemistry analysis (**Figures 2B,C**). As shown in **Figure 2B**, immunoblots of MYOF, EGFR and EPHA2 revealed marked upregulation of these proteins in the membrane fraction, which were consistent with the quantitative proteomic measurements. Higher expression of these proteins was also detected in NPC tissues with metastasis than in those without metastasis.

Bioinformatic Analysis of Differentially Expressed Proteins

To investigate the biological function of the DEPs, the signaling pathways involved were investigated via GSEA analysis. Pathway enrichment analysis showed that the DEPs were primarily involved in biological processes, such as PI3K-AKT pathway, focal adhesions, virus-host interaction, integrin signaling, and metabolic reprogramming (**Figure 3A**), some of which are well known to be associated with metastasis. Of the 294 DEPs, 29 were involved in the Focal Adhesion-PI3K-mTOR (FAK-PI3K-mTOR) signaling pathway (**Supplementary Figure 1**). Intriguingly, all the plasma membrane signaling receptors among the DEPs were involved in this signaling pathway. These plasma membrane receptors, including 3 receptor tyrosine kinases and 7 integrin family members, are compiled in **Table 3**. These findings suggest that the FAK-PI3K-mTOR pathway might play a substantial role in NPC metastasis.

Next, we constructed a protein interaction network using Harmonizome. In total, 731 proteins that interacted with MYOF were acquired from the Harmonizome datasets, among which 65 proteins were present in the DEP dataset (**Supplementary Figure 2**). Interestingly, two of the 10 differentially expressed plasma membrane receptors, EGFR and EPHA2, were among the 65 proteins (**Figure 3B**). An interaction network between the DEPs was assembled also from the Harmonizome datasets (**Supplementary Figure 3**). A partial interaction network focused on the interaction of MYOF, EGFR and EPHA2 is presented in **Figure 3C**. We can observe that only EGFR and EPHA2 have direct interaction with MYOF. We subsequently confirmed the interactions of MYOF with EGFR and EPHA2 using co-immunoprecipitation assay (**Figure 3D**). For further evaluating whether the interactions of MYOF with EGFR and EPHA2 are shared across tumor types, we analyzed MYOF co-expression with the members of common cell membrane receptor kinase (RTK) family from the proteomic data of 375 cancer cell

[1]http://download.baderlab.org

[2]https://amp.pharm.mssm.edu/Harmonizome/

[3]https://www.pathvisio.org/

[4]https://gygi.hms.harvard.edu/publications/ccle.html

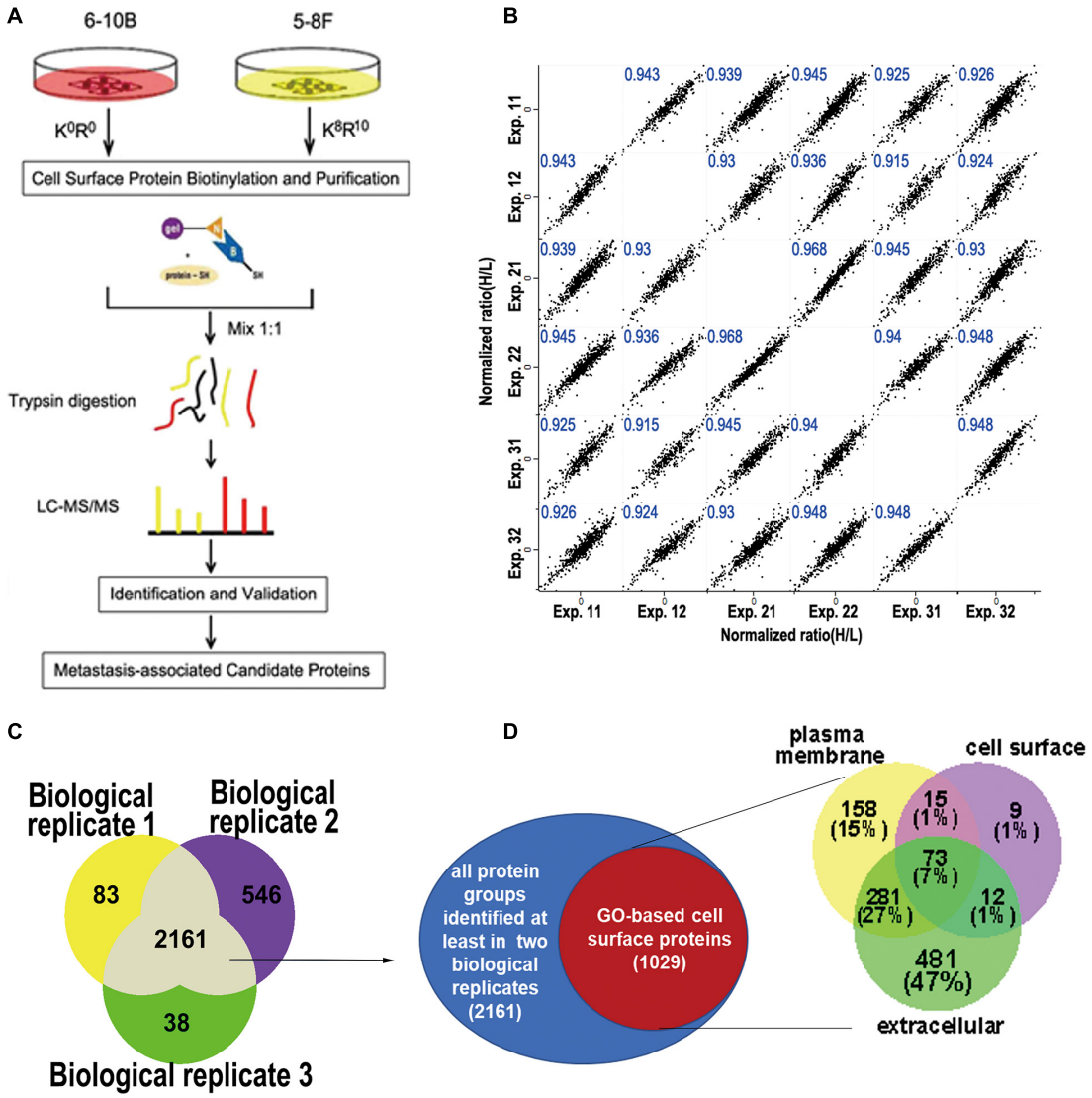


FIGURE 1 | Experimental workflow for cell surface labeling and preparation for mass spectrometry. **(A)** Workflow of SILAC mass spectrometry-based analysis of the cell surface proteome, where 6-10b and 5-8F cells were labeled with "light" and "heavy" amino acids, respectively. **(B)** Scatter plot of protein ratio (5-8F vs. 6-10B) between the biological and technical replicates of the cell surface proteome. The Pearson correlation coefficient is presented in blue at the top of each plot. **(C)** Venn diagram showing the number of identified proteins from three biological replicate experiments. **(D)** Identified proteins grouped by their annotated subcellular localization (UniProtKB). Surface-exposed proteins represented by their annotated detailed categories are shown in the Venn diagram.

lines from 22 lineages in CCLE. It is intriguing that EGFR and EPHA2 are the most related protein of MYOF and the correlation coefficients are 0.68 and 0.61, respectively (**Supplementary Figure 5**).

MYOF Silencing Inhibits the Malignant Phenotype of NPC Cells

To characterize the function of MYOF in NPC, we depleted MYOF by stably transducing the high metastatic 5-8F cells with MYOF shRNA-expressing lentivirus.

The shRNA vectors were constructed and yielded a >50% reduction in MYOF protein levels (**Supplementary Figure 4**). Compared with the control cells, MYOF depletion significantly reduced cellular proliferation rates by 50% (**Figure 4A**). We further investigated whether MYOF knockdown has an impact on migratory and invasive behavior. Our data indicated that the migration of 5-8F cells was severely inhibited in a wound-healing assay (**Figure 4B**). Likewise, the migration of 5-8F cells in a Transwell assay was significantly reduced (**Figure 4C**). In terms of invasion, a nearly 50% reduction in infiltration rate in Transwell invasion assays was observed in MYOF-depleted cells compared with control cells (**Figure 4D**). To extend our findings to tumor tissues, we investigated whether the protein level of MYOF was correlated with metastasis in surgically resected human NPC specimens. IHC was performed on samples

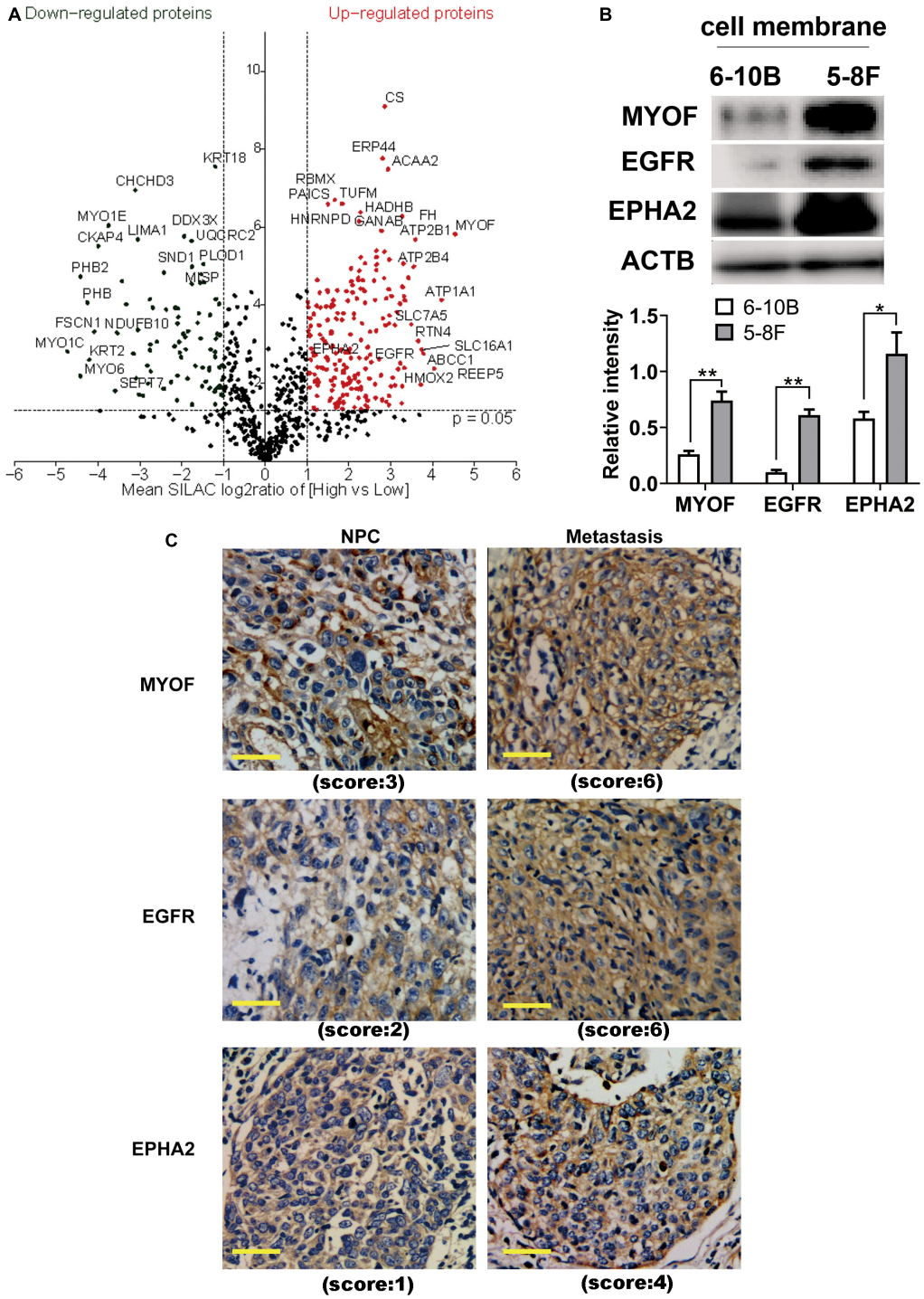


FIGURE 2 | Identification and evaluation of significantly altered proteins between the 6-10B and 5-8F cell lines. **(A)** Volcano plot indicating significantly altered surface-exposed proteins identified in the datasets. Log-transformed *P*-values (*t*-test) against log-transformed fold change in abundance between 6-10B and 5-8F cells. **(B)** Western blot analysis evaluating the expression of proteins that reside in the plasma membrane. The density of the bands was analyzed by using NIH ImageJ software and normalized by the arbitrary units of b-actin. Data are the means ± SDs of 3 experiments. $^{**}p < 0.05$ and $^{**}p < 0.01$. **(C)** Representative immunohistochemistry (IHC) results showing increased detection of the indicated proteins in NPC with metastasis compared with NPC without metastasis.

from 50 cases of stage I–V NPC. The immunohistochemical staining scores indicated that the expression levels of MYOF were significantly higher in NPC with lymph node metastasis than without metastasis ($p = 0.006$, **Table 4**). Likewise, the expression of MYOF was also higher in NPC with distant metastasis than without metastasis ($p = 0.003$, **Table 4**). These results support the notion that the MYOF expression levels are associated with NPC metastasis potential. In summary, these results demonstrated

TABLE 1 | The number of upregulated or downregulated proteins and their subcellular location.

	Total	Subcellular location		
		Cell Surface	Plasma Membrane	Extracellular
	294	46	141	107
Upregulation	206	41	93	72
Downregulation	88	5	48	35

that high expression of MYOF contributes to the malignant phenotype of metastatic NPC cells.

MYOF Knockdown Interferes With Membrane Signaling Receptor Activity and EMT in NPC Cells

In order to clarify the potential mechanism behind the knockdown of MYOF suppressing the malignant phenotype, we investigated whether MYOF knockdown can influence EGFR or EPHA2 activity in NPC cells. As shown in **Figure 5A**, upon EGF stimulation, prominent phosphorylation of the Y1068 residue was observed in MYOF-depleted NPC cells, which led to sustained pEGFR activation. Meanwhile, total EGFR expression levels were increased in comparison with those in vector-transfected cells. These findings demonstrate that MYOF influences EGFR and its phosphorylation-mediated activation status upon EGF stimulation. Previous studies have shown that EGF treatment promotes EMT-like morphological changes in all NPC cell lines (Yip and Seow, 2012). Increasing evidence indicates that EGF activate the EMT process, which promotes cell migration, invasion and metastasis (Wang and Zhou, 2013). Therefore, we speculated that MYOF depletion might influence EMT. To clarify this issue, we examined the expression alteration of EMT markers in MYOF knockdown 5-8F cells upon EGF stimulation. As shown in **Figure 5B**, in the presence of EGF stimulus, blank vector-transfected 5-8F cells exhibited strong induction of vimentin (VIM) expression in parallel with a weak downregulation of E-cadherin (E-cad) at both 24 h and 48 h, whereas the strong induction of VIM was observed at 24 h and 48 h in shMYOF-transfected cells. Moreover, the knockdown of MYOF led to a reduction of EGF-induced VIM, particularly at 48 h. MYOF knockdown also reduced E-cad at both 24 h and 48 h upon EGF stimulus. Likewise, we observed that the knockdown of MYOF led to the increasing of total EPHA2 expression levels and sustained EPHA2 ligand-independent activation, which was evidenced by prominent phosphorylation of the S897 residue. These data suggest that MYOF might interact with the two RTKs, EGFR and EPHA2, thereby regulating their expression and activation in the presence of specific stimulus.

DISCUSSION

Nasopharyngeal carcinoma arises from the nasopharynx epithelium and has high metastatic potential. It has been reported that approximately one third of cases relapse locoregionally or distantly despite intensive, definitive treatment (Cao et al., 2017). Novel treatment strategies beneficial against recurrence or for advanced patients are an important research direction. For example, recently, targeted therapy has demonstrated favorable responses and treatment outcomes (Zhang et al., 2013). Proteomic analysis of differentially expressed cell surface proteins during metastasis has been shown to be an efficient tool for discovering target candidates (Li et al., 2011; Xiao and Chen, 2019).

In present study, we performed the combination of cell surface biotinylation and SILAC-based quantitative proteomics to identify cell surface proteins associated with NPC metastasis. we found that 2161 proteins were present in at least in two biological replicates. Among them, 1029 proteins were identified as cell surface proteins according to GO or UniProt annotation. The number of identified proteins in previous studies aimed at profiling cell surface proteins ranged from less than 400 to approximately 1000 (Lund et al., 2009; Kuhlmann et al., 2018). Therefore, our results are comparable with those of previous studies (Karhemo et al., 2012; Li Y. et al., 2019). To the best of our knowledge, the present study presents the first and largest cell surface proteome of NPC published thus far.

Subsequently, through quantitative proteomic analysis, we identified 294 significantly differentially expressed proteins from the quantified cell surface proteins. Bioinformatic analysis showed that these proteins are involved in multiple pathways, some of which are related to metastasis. For example, in colon cancer, IL13 bind its receptor to activate the FAK-PI3K-mTOR pathway, resulting in cancer metastasis (Bartolomé et al., 2015). Increasing evidence has demonstrated that the mTOR complexes mTORC1 and mTORC2 participate in regulation of cell motility, invasion and cancer metastasis (Zhou and Huang, 2011). In our dataset, the largest number of differentially expressed cell surface proteins were aggregated in this pathway. Moreover, all the differentially expressed cell membrane receptors identified in present study were also involved in this pathway. Based on these observations, it was suggested that the FAK-PI3K-mTOR signaling pathway might be the most important pathway in regulation of NPC metastasis.

Adhesion-related proteins play an important role in the process of metastasis. Some of the differentially expressed proteins, such as Actin-1 (ACTN1), Caveolin-1 (CAV1), and Fibronectin-1 (FN1), were enriched in the Focal adhesion pathway. The roles of these adhesion-related proteins in cancer metastasis are not entirely clear, and some are even controversial. For example, Yu et al. found that CAV1 can promote hepatocellular carcinoma cell progression and metastasis through the Wnt/β-Catenin pathway (Yu et al., 2014), whereas Trimmer et al. demonstrated that CAV1 suppressed tumor growth and metastasis in a murine model of cutaneous squamous cell carcinoma through modulation of MAPK/AP-1 activation (Trimmer et al., 2013). Even in HCC, the function of CAV1 as a tumor suppressor or promoter is still under debate (Yang et al., 2010). These data indicated that malignant cells from different cancer types have different CAV1 expression profiles,

TABLE 2 | Top 20 list of up- and downregulated proteins in 5-8F vs. 6-10B cells.

Symbol	Protein name	Score	Peptides	Log2(MR)	*P*-value	PM	CS	SR	EC
Up-regulated proteins									
MYOF	Myoferlin	323.31	146	4.54	1.50E–06	+			+
ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	323.31	62	4.22	7.34E–05	+			+
ABCC1	Multidrug resistance-associated protein 1	266.82	37	3.78	1.74E–03	+			+
SLC16A1	Monocarboxylate transporter 1	206.15	4	3.74	1.39E–03	+			+
HMOX2	Heme oxygenase 2	323.31	20	3.72	1.09E–02	+			
RTN4	Reticulon-4	285.14	15	3.65	8.20E–04	+			+
ATP2B1	Plasma membrane calcium-transporting ATPase 1	323.31	53	3.59	2.08E–06	+			+
ATP2B4	Plasma membrane calcium-transporting ATPase 4	323.31	24	3.55	1.02E–05	+			
SLC7A5	Large neutral amino acids transporter small subunit 1	247.22	7	3.49	3.11E–04	+			+
SLC1A5	Neutral amino acid transporter B(0)	284.44	15	3.34	7.46E–05	+			+
ATP1B3	Sodium/potassium-transporting ATPase subunit beta-3	107.89	14	3.33	3.99E–03	+			+
PON2	Serum paraoxonase/arylesterase 2	86.501	3	3.29	3.36E–05	+			+
HTRA2	Serine protease HTRA2	26.469	10	3.28	1.18E–02	+			
GNA13	Guanine nucleotide-binding protein subunit alpha-13	26.947	11	3.26	3.27E–02	+			+
EGFR	Epidermal growth factor receptor	323.31	49	3.23	3.04E–03	+	+	+	+
PVR	Poliovirus receptor	165.18	8	3.17	4.38E–03	+	+		+
SLC2A1	Solute carrier family 2	57.723	6	3.15	1.50E–04	+			+
IGF1R	Tyrosine-protein kinase receptor	17.675	11	3.1	1.14E–02	+		+	
GOT2	Aspartate aminotransferase	300.24	20	3.05	7.17E–03	+			+
ICAM1	Intercellular adhesion molecule 1	226.69	19	3.02	5.84E–05	+	+		+
Down-regulated proteins									
MYO1C	Unconventional myosin-Ic	323.31	64	−4.73	1.55E–03	+			+
MYO6	Unconventional myosin-VI	323.31	61	−4.42	6.56E–03	+			+
PHB2	Prohibitin-2	323.31	29	−4.42	1.85E–05		+		+
PHB	Prohibitin	323.31	27	−4.25	8.51E–05	+	+		+
CKAP4	Cytoskeleton-associated protein 4	323.31	45	−3.99	3.06E–06	+			+
SEPT	Septin-7	203.21	17	−3.59	1.60E–02	+			+
SEPT2	Septin-2	323.31	19	−3.43	2.39E–05	+			+
MYO1B	Unconventional myosin-Ib	323.31	41	−3.15	1.73E–03	+			+
LIMA1	LIM domain and actin-binding protein 1	323.31	61	−3.05	2.04E–06	+			
ABCG2	ATP-binding cassette sub-family G member 2	41.72	9	−2.93	2.07E–02	+			
CORO1C	Coronin-1C	167.2	25	−2.76	7.03E–05	+			
IQGAP1	Ras GTPase-activating-like protein IQGAP1	323.31	102	−2.69	1.16E–04	+			+
KTN1	Kinectin	323.31	70	−2.59	8.49E–04	+			
MAP4	Microtubule-associated protein	179.13	18	−2.43	1.39E–02	+			+
ALPP	Alkaline phosphatase	323.31	23	−2.42	1.45E–05	+	+		
SLC25A11	Mitochondrial 2-oxoglutarate/malate carrier protein	111.51	12	−2.37	1.29E–04	+			
SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein	202.11	9	−2.25	9.35E–03	+			+
LMO7	LIM domain only protein 7	323.31	44	−2.16	1.15E–04	+			
MAP7	Ensconsin	21.18	10	−2.11	2.70E–02	+			
CAPRIN1	Caprin-1	106.87	9	−2.1	6.40E–03	+			

The listed proteins were analyzed with multiple biological repeats (heavy high metastatic NPC cell line: light low metastatic NPC cell line). For each protein, the following information is shown: adjusted P-values, logarithm of median ratio (high/low), number of quantified razor peptides plus unique peptides in a replicate. For identification, the peptide and protein FDRs were both set to 1% when using MAXQUANT. Peptides, Razor + unique peptides; MR, median ratio; PM, plasma membrane; CS, cell surface; SR, signaling receptor; EC, extracellular.

which suggests that CAV1 biological functions might depend on the context. In the present study, CAV1 was significantly downregulated in the high metastatic cell line, but its biological role deserves further investigation.

In the list of differentially expressed cell surface proteins, the plasma membrane protein MYOF exhibited the largest fold change. MYOF is a member of the Ferlin family involved in membrane fusion, membrane repair, and membrane trafficking. More and more evidences indicate that MYOF has important significance in clinical diagnosis and targeted cancer therapy (Zhu et al., 2019). However, little is known about its involvement in NPC development and progression. To investigate the function of MYOF in NPC metastasis, we knocked down MYOF in high metastatic 5-8F cells. Consistent with a previous study in

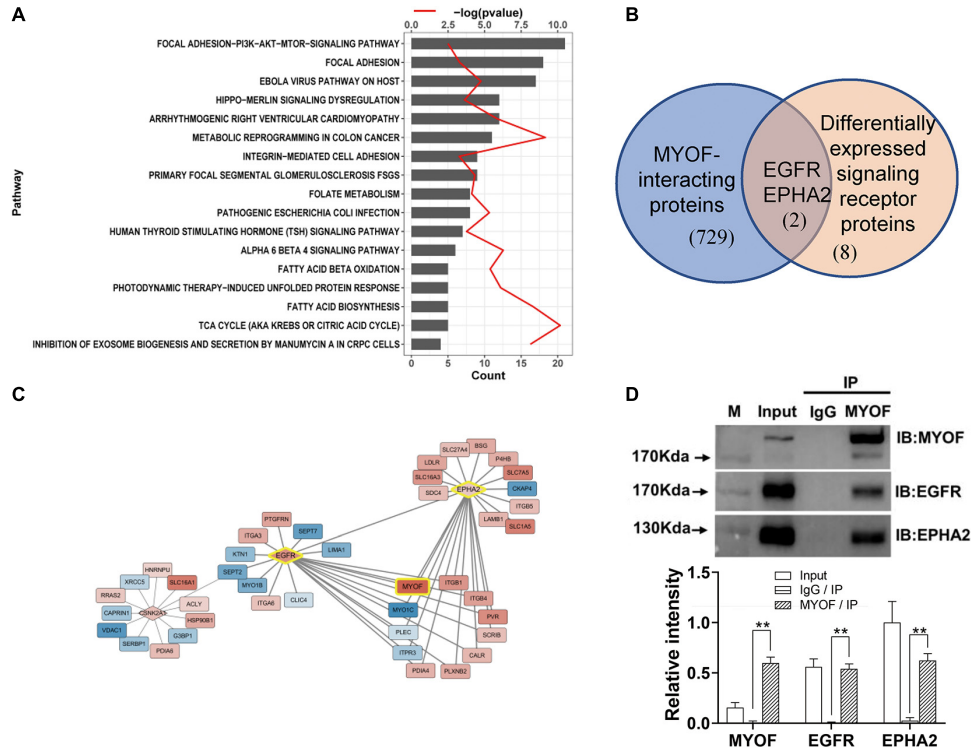


FIGURE 3 | Signaling pathway enrichment analysis and interaction network analysis of DEPs **(A)** Enriched Wikipathway for differentially expressed surface-exposed proteins in 5-8F cell lines vs. 6-10B cell lines. **(B)** Overlapping of MYOF-interacting proteins and differentially expressed signaling receptors in 5-8F vs. 6-10B cells. MYOF-interacting proteins were retrieved using Harmonizome. **(C)** Partial interaction network of differentially expressed proteins focused on the interaction of MYOF, EGFR and EPHA2. **(D)** The interactions of MYOF with EGFR and EPHA2 were confirmed by Co-IP assay. Co-IP using anti-MYOF and negative control antibodies were performed in 5-8F cells. Western blots for all three proteins were performed. The density of the bands was analyzed by using NIH ImageJ software and normalized by the arbitrary units of the band of the input of EphA2. Data are the means ± SDs of 3 experiments. $^{**}p < 0.05$ and $^{**}p < 0.01$.

TABLE 3 | Dysregulated signaling receptors in 5-8F vs. 6-10B cells.

Symbol	Protein name	FAK-PI3K-mTOR	Score	Peptides	Log2(MR)	P-value	PM	CS	SR
EGFR	Epidermal growth factor receptor	+	323.31	49	3.23	3.04E−03	+	+	+
IGF1R	Tyrosine-protein kinase receptor	+	17.675	11	3.1	1.14E−02	+		+
ITGB1	Integrin beta-1	+	310.22	28	2.95	3.91E−05	+	+	+
ITGB4	Integrin beta-4	+	323.31	78	2.6	2.33E−05	+	+	+
ITGA2	Integrin alpha-2	+	323.31	33	2.44	4.13E−02	+	+	+
ITGA3	Integrin alpha-3	+	219.91	19	2.29	1.27E−02	+	+	+
ITGA5	Integrin alpha-5	+	323.31	22	2.14	6.58E−03	+	+	+
ITGB5	Integrin beta-5	+	283.75	18	1.73	1.59E−04	+	+	+
ITGA6	Integrin alpha-6	+	323.31	37	1.61	1.10E−03	+	+	+
EPHA2	Ephrin type-A receptor 2	+	323.31	44	1.45	2.16E−03	+	+	+

breast cancer cells (Turtoi et al., 2013) and Clear-Cell Renal-Cell Carcinoma (Cox et al., 2020), MYOF knockdown markedly *in vitro* inhibit cell proliferation, migration and invasion. Indeed, previous studies demonstrated that MYOF depletion reduced tumor development in a xenograft model of human breast cancer (Turtoi et al., 2013). Moreover, loss of MYOF or pharmacological inhibition of MYOF reduces breast cancer metastasis in an experimental mouse model, which demonstrated that targeting MYOF may be a promising therapeutic strategy in MYOF-driven breast cancer (Zhang et al., 2018). Immunohistochemical analysis showed that MYOF upregulation was related to NPC metastasis. It was in consistent with the finding that MYOF upregulation in many tumors was associated with metastasis (Wang et al., 2013; Blomme et al., 2017; Gu et al., 2020).

Bioinformatic analysis predicted that MYOF can interact with EGFR and EPHA2, both of which were also among the differentially expressed cell surface proteins. There are no experimental evidences supporting the existence of these

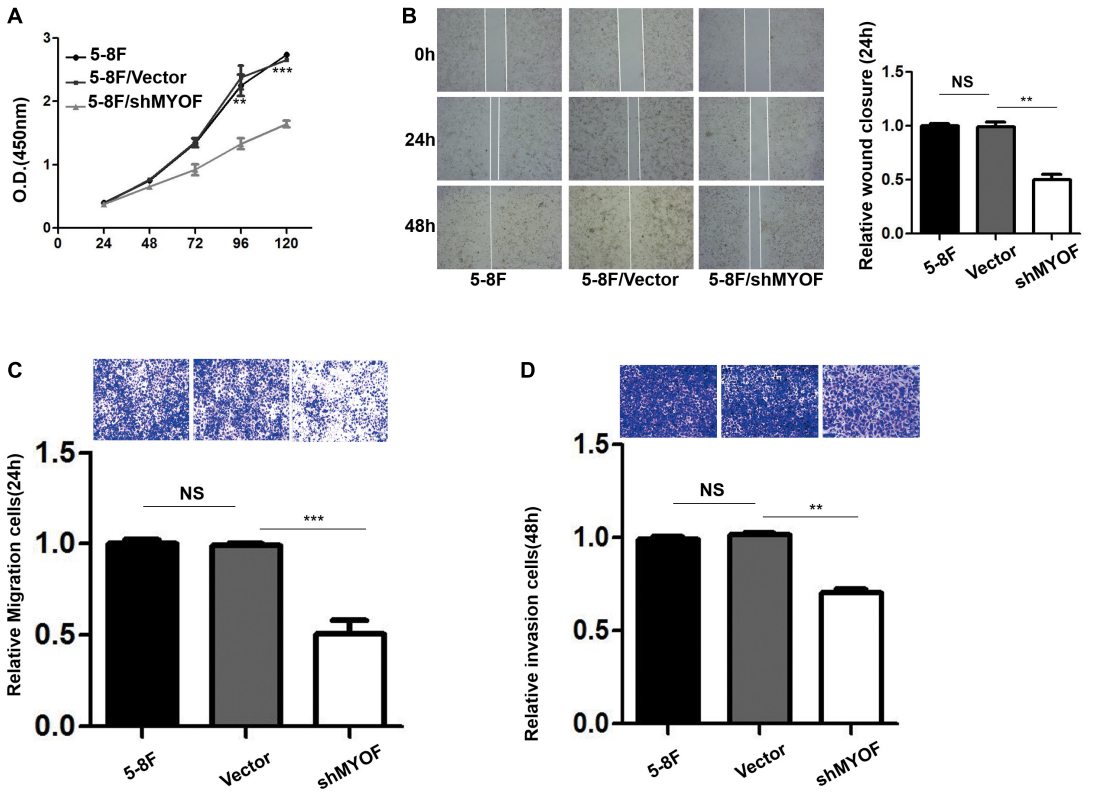


FIGURE 4 | MYOF silencing reduces the malignant phenotype of 5-8F cells. **(A)** Cellular proliferation of control and MYOF knockdown 5-8F cells when passaged for the indicated days. Independent experiments performed in triplicate (**$p < 0.01$ and ***$p < 0.001$, respectively, Students *t*-test). **(B)** Wound-healing assay performed with Vector- or shMYOF-transfected 5-8F cells. Representative images acquired at the indicated time points are shown. The unhealed area was measured in three independent experiments (***p*-value < 0.01). **(C)** Control or MYOF knockdown 5-8F cells were subjected to Transwell migration assays in three independent experiments. Migratory cells were counted under a microscope (**$p < 0.01$). **(D)** Control or MYOF knockdown 5-8F cells were subjected to Transwell invasion assays (with Matrigel). Invasive cells were counted under a microscope (**$p < 0.01$).

TABLE 4 | Association between MYOF expression and clinicopathological characteristics in 50 nasopharyngeal carcinoma cases.

Item	Number	MYOF		P-value
		Low (0-4)	High (5-8)	
Gender				P = 0.374
Male	28	13	15	
Female	22	13	9	
Age(y)				P = 0.441
≥50	27	17	10	
<50	23	12	11	
Primary tumor(T) stages				P = 0.395
T1-2	23	10	13	
T3-4	27	15	12	
Lymph nodes metastasis				P = 0.006**
N0	22	14	8	
N1-N3	28	7	21	
Distant metastasis				P = 0.003**
M0	31	20	11	
M1	19	4	15	

***p < 0.01 determined by χ^2-test, N0 vs. N1-3, and M0 vs.M1.*

interactions. Therefore, we performed Co-IP assay for the first time to verify the interactions of MYOF with EGFR and EPHA2. Based on these observations, we speculated whether the role of MYOF in regulating metastasis is related to its interaction with EGFR or EPHA2. Through shRNA inference assays, we found that MYOF knockdown significantly impaired the expression and phosphorylation level of EGFR, which is consistent with the study of Turtoi et al. in breast cancer cells (Turtoi et al., 2013). EGFR is a well investigated RTK whose activation and signaling contribute to tumor initiation and development, including proliferation, invasion, and metastasis. EGF induced EMT mediates the role of EGFR in promoting metastasis. Our data showed that MYOF knockdown in 5-8F cells significantly inhibited the expression of the EMT markers VIM and E-cad. Therefore, these observations indicated that MYOF knockdown can impair the downstream signaling of EGFR. EGFR is a therapeutic target for multiple types of cancer including NPC (Zhang et al., 2015). However, current EGFR targeted therapies remain somewhat unsatisfactory in the clinic, which might be due to limited knowledge of the mechanism (Peng et al., 2018). An in-depth understanding of MYOF as a regulator of EGFR might contribute to improvement of EGFR targeted therapy.

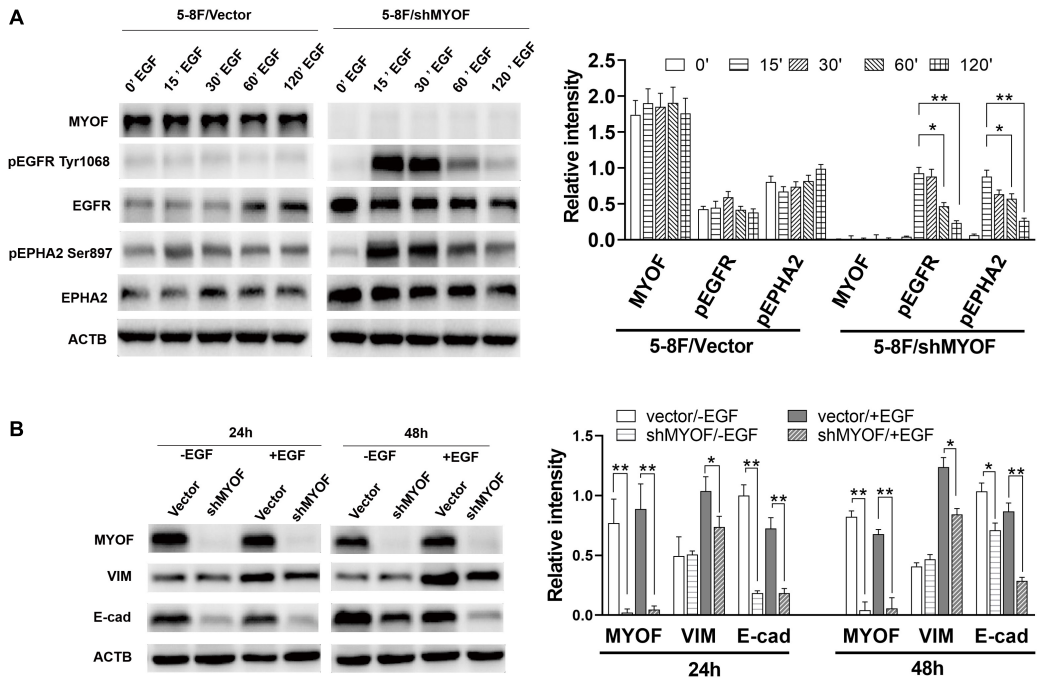


FIGURE 5 | Myoferlin knockdown influences membrane receptor activity and EMT in NPC cells. **(A)** Western blot analysis of time-dependent expression and phosphorylation level of EGFR and EPHA2 following EGF stimulation and myoferlin knockdown. **(B)** MYOF knockdown inhibited EGF-induced VIM expression. The density of the bands was analyzed by using NIH ImageJ software and normalized by the arbitrary units of b-actin. Data are the means ± SDs of 3 experiments. $^{*}p < 0.05$ and $^{**}p < 0.01$.

Ephrin type-A receptor 2 is a member of the Eph RTK family, which plays an essential role in both normal development and disease (Ieguchi and Maru, 2019). Unlike other family members, EPHA2 has a unique feature. It has two activation mechanisms, namely, ligand-dependent and ligand-independent activation, which have distinctive biological effects (Miao et al., 2009). EPHA2 ligand-independent activation, characterized by phosphorylation at S897, has been reported to contribute to metastasis in many tumors, such as colon cancer (Dunne et al., 2016), glioma (Miao et al., 2015), prostate cancer (Tawadros et al., 2012), and NPC (Li J. Y. et al., 2019). Indeed, EphA2 has been evaluated as a drug target using multiple approaches, such as agonist antibodies in leukemia (Charmsaz et al., 2015) and small molecule inhibitors in triple-negative breast cancer (Song et al., 2017). Our results showed that MYOF knockdown has an impact on EPHA2 ligand-independent activation and thus is a regulator of EPHA2 activity. It is the first time reported that MYOF is a regulator of EPHA2 activity. However, the link between MYOF function in metastasis and its regulation of EPHA2 activity deserves further study, which will contribute to the development of new therapy targeting MYOF or improvement of therapies targeting EPHA2. In fact, a previous study supposed that, in addition to EGFR, other RTKs might be regulated by MYOF in various cell types (Turtoi et al., 2013). For example, it was reported that MYOF regulates VEGFR2 stability and function in endothelial cells (Bernatchez et al., 2007). Furthermore, our analysis based on previous pan-cancer proteomic data also demonstrated that EGFR and EPHA2 are the most positive correlated with MYOF across different types. It suggested that MYOF interacts with EGFR and EPHA2 and regulates their activities are universal mechanism by which MYOF exerts its biological functions. In present study, we found that, in addition to EGFR, EPHA2 is a RTK regulated by MYOF in NPC. Our finding verified the supposition and further supported that MYOF might be a potential therapeutic target in NPC.

CONCLUSION

In summary, our study provides a repertoire of cell surface proteins potentially useful for elucidating the mechanisms behind NPC metastasis. Moreover, we demonstrated that MYOF, a regulator of EGFR and EPHA2 activity, might be a potential target for development of new therapies for NPC.

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

ZC and YC designed the experiments and conducted the study supervision. ML, FP, and GW performed the experiments and

analyzed the data. ML, YC, and GW wrote the manuscript. XL helped to analyze the data. MS helped to perform the experiments. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 81372904, 81570537, 81272971, and 81602572), the International Scientific and Technology Cooperation Program of Chia (No. 2015DFA31490), and the Natural Science Foundation of Human province (No. 2017JJ3485).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.621810/full#supplementary-material

REFERENCES

Bartolomé, R. A., García-Palmero, I., Torres, S., López-Lucendo, M., Balyasnikova, I. V., and Casal, J. I. (2015). IL13 receptor alpha2 signaling requires a scaffold protein, FAM120A, to activate the FAK and PI3K pathways in colon cancer metastasis. *Cancer Res.* 75, 2434–2444. doi: 10.1158/0008-5472.can-14-3650

Bernatchez, P. N., Acevedo, L., Fernandez-Hernando, C., Murata, T., Chalouni, C., Kim, J., et al. (2007). Myoferlin regulates vascular endothelial growth factor receptor-2 stability and function. *J. Biol. Chem.* 282, 30745–30753. doi: 10.1074/jbc.m704798200

Blomme, A., Costanza, B., de Tullio, P., Thiry, M., Van Simaeys, G., Boutry, S., et al. (2017). Myoferlin regulates cellular lipid metabolism and promotes metastases in triple-negative breast cancer. *Oncogene* 36, 2116–2130. doi: 10.1038/onc.2016.369

Buehler, U., Schulenburg, K., Yurugi, H., Šolman, M., Abankwa, D., Ulges, A., et al. (2018). Targeting prohibitins at the cell surface prevents Th17-mediated autoimmunity. *EMBO J.* 37:e99429.

Cao, S. M., Yang, Q., Guo, L., Mai, H. Q., Mo, H. Y., Cao, K. J., et al. (2017). Neoadjuvant chemotherapy followed by concurrent chemoradiotherapy versus concurrent chemoradiotherapy alone in locoregionally advanced nasopharyngeal carcinoma: a phase III multicentre randomised controlled trial. *Eur. J. Cancer* 75, 14–23. doi: 10.1016/j.ejca.2016.12.039

Charmsaz, S., Beckett, K., Smith, F. M., Bruedigam, C., Moore, A. S., Al-Ejeh, F., et al. (2015). EphA2 is a therapy target in EphA2-positive leukemias but is not essential for normal hematopoiesis or leukemia. *PLoS One* 10:e130692. doi: 10.1371/journal.pone.0130692

Chua, M., Wee, J., Hui, E. P., and Chan, A. (2016). Nasopharyngeal carcinoma. *Lancet* 387, 1012–1024.

Cox, A., Zhao, C., Tolkach, Y., Nettersheim, D., Schmidt, D., Kristiansen, G., et al. (2020). Knockdown of myoferlin suppresses migration and invasion in clear-cell renal-cell carcinoma. *Anticancer Res.* 40, 3119–3128. doi: 10.21873/anticanres.14293

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511

Dunne, P. D., Dasgupta, S., Blayney, J. K., McArt, D. G., Redmond, K. L., Weir, J. A., et al. (2016). EphA2 expression is a key driver of migration and invasion and a poor prognostic marker in colorectal cancer. *Clin. Cancer Res.* 22, 230–242. doi: 10.1158/1078-0432.ccr-15-0603

Gu, H., Peng, Y., and Chen, Y. (2020). An emerging therapeutic approach by targeting myoferlin (MYOF) for malignant tumors. *Curr. Top. Med. Chem.* 20, 1509–1515. doi: 10.2174/1568026620666200618123436

Ieguchi, K., and Maru, Y. (2019). Roles of EphA1/A2 and ephrin-A1 in cancer. *Cancer Sci.* 110, 841–848. doi: 10.1111/cas.13942

Karhemo, P. R., Ravela, S., Laakso, M., Ritamo, I., Tatti, O., Mäkinen, S., et al. (2012). An optimized isolation of biotinylated cell surface proteins reveals novel players in cancer metastasis. *J. Proteomics* 77, 87–100. doi: 10.1016/j.jprot.2012.07.009

Kuhlmann, L., Cummins, E., Samudio, I., and Kislinger, T. (2018). Cell-surface proteomics for the identification of novel therapeutic targets in cancer. *Expert Rev. Proteomics* 15, 259–275. doi: 10.1080/14789450.2018.1429924

Lai, S. Z., Li, W. F., Chen, L., Luo, W., Chen, Y. Y., Liu, L. Z., et al. (2011). How does intensity-modulated radiotherapy versus conventional two-dimensional radiotherapy influence the treatment results in nasopharyngeal carcinoma patients? *Int. J. Radiat. Oncol. Biol. Phys.* 80, 661–668. doi: 10.1016/j.ijrobp.2010.03.024

Li, G., Xiao, Z., Liu, J., Li, C., Li, F., and Chen, Z. (2011). Cancer: a proteomic disease. *Sci. China Life Sci.* 54, 403–408.

Li, J. Y., Xiao, T., Yi, H. M., Yi, H., Feng, J., Zhu, J. F., et al. (2019). S897 phosphorylation of EphA2 is indispensable for EphA2-dependent nasopharyngeal carcinoma cell invasion, metastasis and stem properties. *Cancer Lett.* 444, 162–174. doi: 10.1016/j.canlet.2018.12.011

Li, M., Peng, F., Li, G., Fu, Y., Huang, Y., Chen, Z., et al. (2016). Proteomic analysis of stromal proteins in different stages of colorectal cancer establishes Tenascin-C as a stromal biomarker for colorectal cancer metastasis. *Oncotarget* 7, 37226–37237. doi: 10.18632/oncotarget.9362

Li, Y., Wang, Y., Mao, J., Yao, Y., Wang, K., Qiao, Q., et al. (2019). Sensitive profiling of cell surface proteome by using an optimized biotinylation method. *J. Proteomics* 196, 33–41. doi: 10.1016/j.jprot.2019.01.015

Lund, R., Leth-Larsen, R., Jensen, O. N., and Ditzel, H. J. (2009). Efficient isolation and quantitative proteomic analysis of cancer cell plasma membrane proteins for identification of metastasis-associated cell surface markers. *J. Proteome Res.* 8, 3078–3090. doi: 10.1021/pr801091k

Mao, Y. P., Xie, F. Y., Liu, L. Z., Sun, Y., Li, L., Tang, L. L., et al. (2009). Re-evaluation of 6th edition of AJCC staging system for nasopharyngeal carcinoma and proposed improvement based on magnetic resonance imaging. *Int. J. Radiat. Oncol. Biol. Phys.* 73, 1326–1334. doi: 10.1016/j.ijrobp.2008.07.062

Miao, H., Gale, N. W., Guo, H., Qian, J., Petty, A., Kaspar, J., et al. (2015). EphA2 promotes infiltrative invasion of glioma stem cells in vivo through cross-talk with Akt and regulates stem cell properties. *Oncogene* 34, 558–567. doi: 10.1038/onc.2013.590

Miao, H., Li, D. Q., Mukherjee, A., Guo, H., Petty, A., Cutter, J., et al. (2009). EphA2 mediates ligand-dependent inhibition and ligand-independent promotion of cell migration and invasion via a reciprocal regulatory loop with Akt. *Cancer Cell* 16, 9–20. doi: 10.1016/j.ccr.2009.04.009

Nusinow, D. P., Szpyt, J., Ghandi, M., Rose, C. M., McDonald, E. R. III, Kalocsay, M., et al. (2020). Quantitative proteomics of the cancer cell line encyclopedia. *Cell* 180, 387–402.

Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006). How many drug targets are there? *Nat. Rev. Drug Discov.* 5, 993–996. doi: 10.1038/nrd2199

Özlü, N., Qureshi, M. H., Toyoda, Y., Renard, B. Y., Mollaoglu, G., Özkan, N. E., et al. (2015). Quantitative comparison of a human cancer cell surface proteome between interphase and mitosis. *EMBO J.* 34, 251–265. doi: 10.15252/embj.201385162

Peng, H., Tang, L. L., Liu, X., Chen, L., Li, W. F., Mao, Y. P., et al. (2018). Anti-EGFR targeted therapy delivered before versus during radiotherapy in locoregionally advanced nasopharyngeal carcinoma: a big-data, intelligence platform-based analysis. *BMC Cancer* 18:323. doi: 10.1186/s12885-018-4268-y

Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., et al. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* 2016:baw100. doi: 10.1093/database/baw100

Song, W., Hwang, Y., Youngblood, V. M., Cook, R. S., Balko, J. M., Chen, J., et al. (2017). Targeting EphA2 impairs cell cycle progression and growth of basal-like/triple-negative breast cancers. *Oncogene* 36, 5620–5630. doi: 10.1038/onc.2017.170

Sun, Y., Yi, H., Zhang, P. F., Li, M. Y., Li, C., Li, F., et al. (2007). Identification of differential proteins in nasopharyngeal carcinoma cells with p53 silence by proteome analysis. *FEBS Lett.* 581, 131–139. doi: 10.1016/j.febslet.2006.12.008

Tawadros, T., Brown, M. D., Hart, C. A., and Clarke, N. W. (2012). Ligand-independent activation of EphA2 by arachidonic acid induces metastasis-like behaviour in prostate cancer cells. *Br. J. Cancer* 107, 1737–1744. doi: 10.1038/bjc.2012.457

Trimmer, C., Bonuccelli, G., Katiyar, S., Sotgia, F., Pestell, R. G., Lisanti, M. P., et al. (2013). Cav1 suppresses tumor growth and metastasis in a murine model of cutaneous SCC through modulation of MAPK/AP-1 activation. *Am. J. Pathol.* 182, 992–1004. doi: 10.1016/j.ajpath.2012.11.008

Turtoi, A., Blomme, A., Bellahcène, A., Gilles, C., Hennequière, V., Peixoto, P., et al. (2013). Myoferlin is a key regulator of EGFR activity in breast cancer. *Cancer Res.* 73, 5438–5448. doi: 10.1158/0008-5472.can-13-1142

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., et al. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13, 731–740. doi: 10.1038/nmeth.3901

Vuckovic, D., Dagley, L. F., Purcell, A. W., and Emili, A. (2013). Membrane proteomics by high performance liquid chromatography-tandem mass spectrometry: analytical approaches and challenges. *Proteomics* 13, 404–423. doi: 10.1002/pmic.201200340

Wang, W. S., Liu, X. H., Liu, L. X., Lou, W. H., Jin, D. Y., Yang, P. Y., et al. (2013). iTRAQ-based quantitative proteomics reveals myoferlin as a novel prognostic predictor in pancreatic adenocarcinoma. *J. Proteomics* 91, 453–465. doi: 10.1016/j.jprot.2013.06.032

Wang, Y., and Zhou, B. P. (2013). Epithelial-mesenchymal transition—a hallmark of breast cancer metastasis. *Cancer Hallm.* 1, 38–49. doi: 10.1166/ch.2013.1004

Xiao, Z., and Chen, Z. (2019). Deciphering nasopharyngeal carcinoma pathogenesis via proteomics. *Expert Rev. Proteomics* 16, 475–485. doi: 10.1080/14789450.2019.1615891

Xiao, Z., Li, G., Chen, Y., Li, M., Peng, F., Li, C., et al. (2010). Quantitative proteomic analysis of formalin-fixed and paraffin-embedded nasopharyngeal carcinoma using iTRAQ labeling, two-dimensional liquid chromatography, and tandem mass spectrometry. *J. Histochem. Cytochem.* 58, 517–527. doi: 10.1369/jhc.2010.955526

Yang, S. F., Yang, J. Y., Huang, C. H., Wang, S. N., Lu, C. P., Tsai, C. J., et al. (2010). Increased caveolin-1 expression associated with prolonged overall survival rate in hepatocellular carcinoma. *Pathology* 42, 438–445. doi: 10.3109/00313025.2010.494293

Yip, W. K., and Seow, H. F. (2012). Activation of phosphatidylinositol 3-kinase/Akt signaling by EGF downregulates membranous E-cadherin and beta-catenin and enhances invasion in nasopharyngeal carcinoma cells. *Cancer Lett.* 318, 162–172. doi: 10.1016/j.canlet.2011.12.018

Yu, H., Shen, H., Zhang, Y., Zhong, F., Liu, Y., Qin, L., et al. (2014). CAV1 promotes HCC cell progression and metastasis through Wnt/beta-catenin pathway. *PLoS One* 9:e106451. doi: 10.1371/journal.pone.0106451

Zhang, J. W., Qin, T., Hong, S. D., Zhang, J., Fang, W. F., Zhao, Y. Y., et al. (2015). Multiple oncogenic mutations related to targeted therapy in nasopharyngeal carcinoma. *Chin. J. Cancer* 34, 177–183.

Zhang, L., Chen, Q. Y., Liu, H., Tang, L. Q., and Mai, H. Q. (2013). Emerging treatment options for nasopharyngeal carcinoma. *Drug Des. Devel. Ther.* 7, 37–52. doi: 10.2147/dddt.s30753

Zhang, T., Li, J., He, Y., Yang, F., Hao, Y., Jin, W., et al. (2018). A small molecule targeting myoferlin exerts promising anti-tumor effects on breast cancer. *Nat. Commun.* 9:3726. doi: 10.1038/s41467-018-06179-0

Zhou, H., and Huang, S. (2011). Role of mTOR signaling in tumor cell motility, invasion and metastasis. *Curr. Protein Pept. Sci.* 12, 30–42. doi: 10.2174/138920311795659407

Zhu, W., Zhou, B., Zhao, C., Ba, Z., Xu, H., Yan, X., et al. (2019). Myoferlin, a multifunctional protein in normal cells, has novel and key roles in various cancers. *J. Cell. Mol. Med.* 23, 7180–7189. doi: 10.1111/jcmm.14648

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 Li, Peng, Wang, Liang, Shao, Chen and Chen. 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